Does room ambience have advantages over artificial post processed vocal ambience for recorded lead vocals in lowdynamic music
A preliminary study

Dan Nyberg

Luleå University of Technology
Bachelor thesis
Audio Technology
Arena, Media, Music and Technology
Division of media and adventure management
Does room ambience have advantages over artificial post processed vocal ambience for recorded lead vocals in low-dynamic music: A preliminary study

Dan Nyberg

Arena Media Music and Technology
Luleå University of Technology
The department of music and media
-2007-

Supervisor: Nyssim Lefford
Sonic Studio, Interactive Institute, Piteå
1. Abstract

This senior thesis examines if there are any advantages to using room ambience over post processed vocal ambience for recorded lead vocals in low-dynamic music. The thesis contains a literature review, an analysis of the techniques of several recording engineers and a listening test. The literature review examines the use of effect on vocals in music. The technique analysis makes use of published interviews from the book, *The mixing engineers handbook* by B. Owsinski, examine the common practice of mixing low-dynamic and is based on the interviews. The results from this analysis produced the mixing constraints for the mixing of the stimuli’s for the listening test. The listening test evaluates five different environments in which vocals have been recorded. The five subjects in the listening test have all a technical audio education. All environments have different acoustical characteristics and include: a church, a song booth, an entrance hall to an apartment, an anechoic chamber and a post processed song booth with ambience effects. The results from the listing test show indications on how humans perceive room size and offer insight into the positive and negative affects of the perceived room size, the amount of volume, intelligibility, reflections and which one of the environment are most and least preferred and why. The conclusion of this thesis is that preliminary findings suggest that there could be an advantage of using room ambience over post processed ambience on lead vocals.
Thanks to;
My supervisor Nyssim Lefford for all support and comments under this project, Tobias Norlén, Katarina Nilsson for assisting under the recordings, Anders Bolin for the vocals, The musicians for the music, Andreas Risberg for programming the listening test interface, Roger Johnsson for helping with the acoustical measurements, and last my family for supporting me.
Table of Content

Abstract

1. Introduction 5

2. Hypothesis/research question 5-6

3. Methodology 6-9

4. Background 9-14
   4.1 What do effects do for vocals? 9-14
      4.1.1 Equalizer 9-10
      4.1.2 Compression 10
      4.1.3 Spatial imaging 10
      4.1.4 Different types of reverbs 11-12
      4.1.5 Delay 12-14

5. Post processing vocals; interview analysis 14-33
   5.1 Analysis of the interviews 15-25
   5.2 Conclusions of the interview analysis 26-29
   5.3 Interesting findings of the interviews 29-31
   5.4 Discussion on the interview analysis 31-33

6. Mixing constraints of the vocal stimuli’s and musical arrangement 33-35

7. Listening test 36
   7.1 Listening test implementations 37-38
   7.2 Evaluation of the stimuli’s in the listening test 38
   7.3 Expected data from the listening test 39
   7.4 Data from the listening test 39-47
   7.5 Conclusions from the listening test 48-51
   7.6 Analysis of the listening test 51-56
   7.7 Discussion of the listening test 56-57
   7.8 Interesting findings of the listening test 57-58

8. Discussion on the work, methodology and future work 58-60

9. Conclusions of the work 60-62

10. Reference 63

11. Appendix 64-76
1. Introduction

When recording music, post processing of the vocals with ambience effects is a standard for placing the vocals in the mix, this includes the spatial placement horizontality and in the perceived depth of the audio field. Placement of the vocals in the horizontal plane is often achieved by panning the sound left or right. This is done by changing the volume of the sound in each speaker. The depth is often created by changing the volume of the vocals or the use of ambience effects such as reverb. The ambience effects can also blend the vocal sound into the arrangement and sound as a the instruments and vocals are in the same environment; this is done by adding artificial room ambiance, or less reverb. When the desired amount of ambiance is added then the environment fits the vocals to the rest of the arrangement.

For making the vocal stand out more from the arrangement additional delay can be added, this effect, can give arise to a variety of different characteristics to the vocals, ranging from doubling effect with a short delay time, a bigger sound with a shorter delay time and with a longer delay time create echo effects. These delay times are imitations of real acoustical phenomenon which occur when doubling a sound or hearing an echo in nature [1]. The delay time is one parameter that humans relate to the spaciousness of the sound is. The level of the delay and the delay time is in relationship with each other, with a greater/longer delay time the more spacious the sound appears and the length of the delay time is perceived easily. The opposite happens when a short delay time is used. Then the volume level of the delay has too be increased [3]. A more detailed explanation on what delay times causes what effect are stated in the background section.

It is no yet clear whether, post processing of the vocals with ambiance effects are necessary to place the vocals in the mix is more effective (easier to make the vocal in balance with the mix) than using the real room acoustics where the vocals is recorded? And if so, what types of rooms are most effective? This could make mixing vocals easier for the engineer, only buy choosing the right environment or procedure when recording are being made. For the purposes of this study, this proposed question is broken down in to two smaller questions:

- How are mixing engineers working with effects with the emphasis on ambience effects in low-dynamic music?
- How is the perception of the vocals changed with different environments?

2. Hypothesis/research question

In dynamic music, such as rock ballads and softer pop music, where the instruments are sounding low in volume, vocals tend to be the focus. If the vocals are poorly preformed or the vocal sounds artificial (it does not attain the same ambiance in the sound as the instruments because it was recorded in another studio with different acoustics) in the mix, then the song is usually considered unpleasant. Dynamic music gives the vocals more room in the sound image and the vocals are perceived as better. In this case the post processing of the vocal with ambience effects has an advantage, due to the precise control over the parameters that the reverb attains. If the vocal do not attain the desired ambiance from the recording studio to match the other instruments, then the mixing

---

[1] The lead vocals placement is always in the centre of the sound image (the perceived placement between the speakers in height, left to right and in volume.) and tends to be the focus of the song. The placement of the ambience effects together with the vocals are often added subtly for creating extra personality to the sound and making the vocals more interesting to listen to. [1]
engineer can modify the ambiance effect to match the recorded ambience from the instruments and thereby placing the vocals in unity with the instruments.

In low-dynamic music however, such as Rock and harder pop music, where the instruments are taking up more room in the sound image and are often mixed more dry (not using an abundance of effects in the mix) then the vocals will have many more sound sources too fight with, too achieve intelligibility and clarity, frequency- and volume-wise, in the sound image. For example, such as frequency overlaps between instrument and vocal range. A distorted guitar, in the critical frequency range of the vocal can make the vocal less intelligible (tydig) or with a too low volume level on the vocals can make the vocal hard too hear in the arrangement, and a room with insufficient acoustics can colorized the sound with unwanted reflection [2]. This can make the vocal sound including the natural ambience which follows the recording to be perceived not as clearly as in the perception of the vocals in the dynamic music.

This suggests that the natural ambience can effectively be used to place the vocals in the sound image without the use of post processing with ambiance effects on vocals in music. And by getting a vocal environment that correlates with the instruments’ environment equally as post processed vocal with ambiance effects correlates with the instruments’ environment, then the step of post processing the vocals with ambiance effects can be eliminated.

3. Methodology

To explore this problem space an experiment has been conducted in which the vocals to a low-dynamic song have been recorded in four different environments along with one post processed vocal, all the environments have different acoustical characteristics. These are mixed as stimuli for a listening test to distinguish how each one of the environments are perceived by listeners, and if all the subjects prefer one environment over another.

The post processed vocal is mixed with mixing constraints that are based on the results of an interview analysis which will be discussed in detail in the section, mixing constraints.

The vocals are added to the low-dynamic music arrangement which is recorded with multi-track technique one instrument at a time. The vocals have no over dubbing, only one track with male vocal.

The post processed vocal is mixed with an ambience effect that is mixed following conventional practice for vocals in low-dynamic rock music.

For further enhancement of the comparison, the background instruments will have the same volume-, pan- and frequency-mix settings in all stimuli. These criteria are stated in the mixing constraints section. One instrument background mix will be made. Only the vocal will be different in each one of the stimuli’s. The vocals have a consistent volume and placement in all of the stimuli. This is achieved by using a constant distance between the singer and microphone:

- 6 cm between the pop eliminator and the microphone
- 21 cm between the pop eliminator and the singer
- The same microphone with the same adjustments in all recordings
- The same microphone preamplifier with the same adjustments in all recordings.

The adjustments and characteristics of the microphone and pre amplifier can be seen in table 7 in the appendix.

In the four environments which the vocals are recorded, measurements of the decay times have been made. The measurements include decay time on 9 frequency bands, ranging from 31,5Hz to 8 kHz in third bands, and the time for the first reflection to reach the microphone (pre delay time) are measured. The sound source in the decay time measurements was 20 kHz pink noise feed...
through a monitor speaker with a limited frequency range of 75 Hz – 18 kHz and an octave band analyser.

In the following description of the environments only one decay time and one angle of the first reflection are stated. For a full view of measurements on decay time and reflection/pre delay time from other positions angles and frequency bands see tables 8-11 in the appendix.

- **Church**
  - Dimensions, width, length, height: 14m*15m*10m
  - Material in walls, floor, ceiling; plywood; wooden panel, wooden floor, wooden panel ceiling.
  - Microphone position; 5m from the back wall, 7m from the side walls.
  - Decay time; from position a) at 1 kHz is 0.974 seconds
  - Reflection/pre delay time; at -45° from sound source, 0.011 seconds

- **Vocal booth (K9)**
  - Dimensions, width, length, height; 2.95m, 2.06m*1.77m, 2.06m*2.36m
  - Material in walls, floor, ceiling; left wall; absorbent, right wall; wooden panel, back wall; wooden panel, front wall; mineral wool, ceiling; absorbents, floor; parquet flooring.
  - Microphone position; 0.88m from front wall window.
  - Decay time; from position a) at 1 kHz is 0.326 seconds
  - Reflection/pre delay time; at -45° from sound source, 0.006 seconds

![Figure 1.1](image)

*Figure 1.1: a) Recording and first measurement microphone position, b) second measurement position c) third measurement position.*
Figure 1.2:  a) Recording and First measurement microphone position, b) second measurement position.

- Entrance Hallway in a apartment
  - Dimensions, width, width, length, height; 1m, 1.6m*2.6m*2.45m
  - Material in walls, floor, ceiling; left wall; drywall, right wall; masonite closet doors, floor; parquet flooring, ceiling; concrete.
  - Microphone position; 1.6m from the entrance of the apartment.
  - Decay time; from position a) at 1 kHz is 0.423 seconds
  - Reflection/pre delay time; at -45° from sound source, 0.011 seconds

Figure 1.3; a) recording microphone position, b) measurement position.

- Anechoic environment
  - The design has no reverberation time at all.

The post processed vocal has the vocal from the room design for vocal (K9), and a plate reverb added. The plate-reverb that is used are the D-verb is a reverb plug-in which follows with the Pro tools recording software, the algorithm used for this work is large plate with variables parameters, what these settings are can be seen table 1 in section; Mixing constraints.

The choice of the plate reverb resulted from the analysis of interviews done by B.Owsinski in his book The Mixing engineer’s handbook. The analysis provided a basis for determining common practice techniques in mixing vocals in rock music and subsequently a basis for the experimental methodology. The analysis is also used for finding a consistent way of mixing the instruments
making the vocals in focus, but not too low, so the instruments are perceived low in volume versus the vocal volume.

The Mixing constraints are based on the results drawn from the interview analysis, and implemented in the mix of the all the stimuli’s.

The four environments and the post processed stimuli are evaluated using a qualitative approach and the listening test. The goal is to understand and give explanations for how the environments are perceived and which ones are most and least preferred. It includes an evaluation of different parameters of the vocal environments such as amount of reverberation time, volume, and intelligibility (tydlighet) of the lyrics.

4. Background:

4.1 What do effects do for vocals?
This section describes the effects used when mixing low-dynamic music and how each effect has an influence on the sound, why it is relevant in mixing music. The emphasis is on reverb and spatial imaging. Section; Post processing vocals and mixing reference stimuli: Interview analysis, examines each effect, closely from the engineer’s point of view when they are applied to vocals.

4.1.1 Equalizer
This section is to illustrate how equalizers work and clarify generally how different frequency ranges corresponds to different sounds and also, what frequencies can be cut or boosted when mixing vocals and instruments and why.

The equalizer is a volume changer, which works on specific frequencies of a sound. It can boost a particular frequency or cut a frequency.

The frequency range can be divided into six ranges for an overview.

<table>
<thead>
<tr>
<th>Frequency Range</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hi highs</td>
<td>&gt;8000 Hz</td>
<td></td>
</tr>
<tr>
<td>Highs</td>
<td>5000-8000 Hz</td>
<td></td>
</tr>
<tr>
<td>Midranges</td>
<td>800-5000 Hz</td>
<td></td>
</tr>
<tr>
<td>Ooozone</td>
<td>200-800 Hz</td>
<td></td>
</tr>
<tr>
<td>Bass</td>
<td>40-200 Hz</td>
<td></td>
</tr>
<tr>
<td>Low Bass</td>
<td>&lt;40 Hz</td>
<td></td>
</tr>
</tbody>
</table>

[1] Low bass is also often called the “Sub bass” and are found in low bass of kick drums and bass guitars. At this frequencies people have a hard time discern pitch. [1]

Bass, this is the rage often boosted when using the bass tone control on a home stereo. [1]

Ooozone, this rage gives the sound an extremely unclear and muddy characteristic if it is added too much, and can cause fatigue. [1]

Midranges, humans are particular sensitive to this frequency range; a boost of 1 dB in this rage can be perceived as a boost of 3 dB in another frequency region. In this rage the human speech is centred. Cutting frequencies in this rage on vocals can give arise to an unnatural sound, because humans are hypersensitive in this frequency region. [1]
Highs, this rage are the one boosted on the treble tone knob in a home stereo, this is often boosted in mastering of a record too make things brighter and more present. [1]

Hi-Highs, in this rage cymbals and higher harmonics are found, by adding a little volume to this rage on certain instruments can make it sound like a higher quality recording, but too much can make it irritating. [1]

4.1.2 Compression
This section explains how compression effects work, why they are used in mixing, and also how to control the sound to avoid volume collisions between vocals and instruments.

The compressor has two major functions: to get a better signal-to-noise ratio and to stabilize the sound image between the speakers, which results in more presences. A dynamic piece of music contains both loud and soft sounding sounds. High volume sections must be turned down so that the signal doesn’t become distorted in the playback signal chain. Turning down the over all level makes the soft sounding sounds becomes ever softer and this gives arise to a bad signal-to-noise ratio. This is particularly problematic if the recordings are made on tape machines [1] or using a low bit rate.
Using a compressor during the recording session makes the loud sound softer and the softer sound louder which give a better signal-to-noise ratio. [1]

To stabilize the image of the sound and give it more presence, the compressor can be used. A sound naturally bounces up and down in volume, and when combining different sounds that naturally bounce up and down the result becomes very erratic. Using a compressor on the sound stabilizes it and smoothes it out the movements of the sounds that result from these moment-to-moment fluctuations in volume. This makes the sound not so erratic and the listener can focus in on it better. Therefore the sound seems clearer and more present in a mix. Once the sound has been compressed, stabilized, then the volume can be boosted. [1]

The compressor can make the attack of the sound more sharper, when compressing the loud signal then the volume becomes lower and reaches the maximum much quicker and with a shorter and sharper attack. That makes the mix sound tighter, punchier, more distinct and more precise. With a higher quality, fast compressor can actually help to remove “spikes” on the attack of the sound, softening the sound. [1]

4.1.3 Spatial imaging
This section is to get an understanding in the sonic parameters that impact the placement of the sound in the sound image and to establish an understanding of how the different instruments and vocal can be placed in the sound image, horizontally and depth.

When placing a sound in the stereo field then there are three basic parameters of sound corresponding to the X, Y and Z axes. By panning the different sound left, right or centre then the sound will appear corresponding in the image left, right or centre, panning correlates to the X axis, the sound that are louder in one ear appears to come from that position in space. Sounds that are louder appear closer the listener and sound that are softer appears to the listener at a more distant. To place a sound in the background, one thing too do is to lower the volume of that sound and by doing the opposite makes the sound feel closer to the listener. Volume correlates to the Z axis. On the Y axis frequencies determines the placement of the sound. The higher the frequencies are the more they tend to be in the upper part of the sound image and the opposite happens with the lower frequencies they tent to be at the bottom of the sound image. [4]
4.1.4 Different types of Reverbs

These following sections are the focus on ambience and understanding the functions of the reverb. This sets a basis for describing the common usage of the reverbs.

Natural reverb

Reverb is hundreds of hundreds of delays, when a sound is created it travels through the room with about a velocity of 770 miles per hour; this sound reflects on the walls, ceiling and floor and comes back as hundreds of different delay times. When combining all these delay times together it is called reverb [1]. The natural reverberation starts with first reflections that arrives from the roof and walls, in the centre of a concert hall with a substantial delay (often called pre delay). These reflections is followed by progressively shorter reflections, in a vast numbers, from all directions, were the sound reflects, until they trail away beyond the threshold of hearing. In music rooms or concert halls, the frequencies in each reflection are absorbed more or less rapidly then others, this is called coloration. This occurs both in natural and artificial reverberation, the build up and blending of the natural reverberant sound depends on the three dimensional system of reflections and decay which enriches the quality at just that stage when coloration is likely be most severe. However the artificial reverberation differ from the natural acoustics. Its use is more acceptable if the coloration and density of the decay products from the artificial reverberation is perceived as musical in quality. [4]

Variable Reverb

In this section the reverb will be explained in two sub sections: what parameter goes in a variable reverb unit and two types of reverb units.
The two units are echo chamber and plate reverb: They are the primary reverb units used by the engineers in the interview analysis. (How the reverbs that are used on vocals can be seen in the topic; Post processing vocals and mixing reference stimuli: Interview analysis.)

Changeable parameters of a variable reverb unit

The most important factor in an enhanced, natural-sounding acoustic is the reverberation time, and by using electronics, this could be set to a variety of values in several frequency ranges.

For example: depending on the period the composition was written and the style, classical music generally sounds best in a hall with a midrange decay of 1, 8-2,5s. However, the low frequencies may require longer decay time say 2, 5-3, 0 s, and it will be necessary to have a cross over frequency, at 600-800 Hz. Treble decay may need too have a shorter decay time with a crossover at 6 kHz or higher. [4]

The length of decay in one or more bands might be dependent on the presence or absence of gaps in the sound, so that it is relatively short when the sound is continuous (for vocals or instrument clarity) but longer in pauses. This modification is to be used cautiously. It may muddy and confuse the general effect. On lead vocals decay time helps to define phrasing, but it is also to supporting lines that have gaps in different places in the song. [4]

Reverb fills out the space between the speakers, and masks other sounds in the mix because in a digital reverb all the delays are panned to nearly hundreds of different places in the sound image [1].

A variable reverb unit has certain parameters that control the reverb.

- **Room Types:** This parameter is to simulate different types of rooms with pre delay time, diffusion and so fourth. The parameters can be chanced separately.
• **Reverb time**: Reverb time can also be changed and it is the duration and length of the reverb that is altered.

• **Pre delay time**: The pre delay time is the time it takes for the direct sound to reflect back. All rooms have a pre delay time different from each other. A medium-sized room has about 30 ms of pre delay time and a coliseum can have about 100 ms of pre delay. This time in reverb units can be altered.

• **Diffusion**: Diffusion is the density of the echoes in the reverb. With a low diffusion a smaller amount of echoes will occur and every echo is easier to locate in the sound. The opposite will happen with the use of high diffusion. High diffusion uses more echoes more tightly spaced together and gives more diffusion to the sound. High diffusion has a sweeter and smoother characteristic, when low diffusion is more intense. High diffusion is often more preferred for vocals.

**Echo chamber**;
Echo chambers were the first way of enhancing reverberation. This technique had the advantage of using real acoustics but these chambers were often too low in volume. On one side of the chamber an omni directional loudspeaker was placed so that no direct sound would hit the microphone placed on the other side of the room. The sound would then travel through the chamber from the loudspeaker, reflecting the sound on the walls, and be picked up by the microphone. Different kinds of room shapes were constructed in the chambers too make the sound travel further, so the reverberation time would be longer, but ones laid out, then the reverberation time would be fixed. [4]

**Plate reverb**;
The reverberation plate has one transducer that delivers the sound to the plate and brings it to vibrate and two other transducers that acts as microphones and picks up the vibration. The contact microphones are placed asymmetrical on the plate too not pick up any common set of vibration antinodes of the plate. The use of the two contact microphones simultaneously, is often when stereo is needed.
The plate is a thin steel sheet suspended in tension at each corner from a steel frame, the sheet has a minimum size of $2 \times m^2$ with a maximum thickness of 0.5 mm. This is to reduce the metallic quality of the resonances to proportions that is acceptable for most purposes. The thickness gives the plate good transverse vibrational properties.
To control the reverberation time, a thin stiff, porous foil is rigidly held at a controlled but variable distance from the plate. This is also for damping the frequency characteristics of the plate to simulate the high frequency absorption of the air in a moderate sized room.
The distance between the damping and the plate determines the reverberation time. With the broadest distance, 120 mm, the reverberation time is 1.5 s at 10,000 Hz and 5.3 s at 500 Hz. At the narrowest distance of 3 mm the reverberation time is 0.3 s. The damping is either controlled mechanically or motorized.
The plate has a limited direct path between the sending transducer and the contact microphones, so for long reverberation times, extra delay needs to be required. [4]

### 4.1.5 Delay
This section clarifies what delay is and what different delay times create certain characteristics in vocals or instruments, such as doubling, fattening or echo too the sound.
Delay time, over 100ms the delay time are referred as “echo”, however, not “echo” in laymen can mean “reverb”. For this purpose “echo” is a delay time greater then 100ms and not reverb.

Delay time has a relationship with distance;

Example: If two microphones are placed at two different distances from a sound source the sound will then reach the closer microphone first before it reaches the second one, this creates a delay. With at longer distance between these two microphones the longer the delay will be, and as the distance decreases the smaller the delay will be. Delay will arise when using more the one microphone on a sound and is especially helpful when recording drums. The overhead microphones above the drum set will create a corresponding delay time between the overhead microphones and the snare microphone. Besides delay time, the phase cancellation must be considered which occur with very short delay times. [1]

>100ms delay

When using a delay time over 100ms it is important too fit the delay time into the tempo of the song, otherwise it will throw off the timing of the song. The delay time should be in time, a multiple of, or an exact fraction of the tempo. [1]

Calculation of delay time to the tempo of a song;

\[
\text{Delay in ms} = \frac{60000}{\text{bpm}} \times \frac{1}{4}\text{note}
\]

[5]

By taking the answer and dividing it by two the lower denominations are known or multiplying the lower denominations or quarter note delay time value by 1,5 the dotted values are known or multiplying them with 667 the triplet value is known [5]. A 100ms long delay creates a dreamy effect and are commonly placed in slower tempos where there is room for additional sounds, in quicker tempos with more notes and instruments the more sparsely the use of delay becomes.

[1]

60 to 100ms

This delay time is audible and is commonly referred to as “slap” as on the vocals of Elvis Presley and in rockabilly music. This effect is commonly used on thin and irritating vocals or sounds to make them seem fuller. It can help obscure bad vocal techniques or pitch problems. It can bury any bad sound, but too much “slap” makes the mix sound bad. “Slap” can make the vocal seem less personal. If the singer is good then the delay might not be needed. [1]

30 to 60ms

The use of this delay time is often called “doubling” It makes one voice or instrument, sound like two or double tracked. When a part is sung or played twice the natural delay time between the takes is about 30 to 60 ms, and when adding a delay with this time, then the sound feels like it has been played twice. It can also obscure, as the “slap” delay, any bad performance or sound and bury it in the mix. However, it takes away the purity and clarity of the sound when it is added too much. [1]
This delay time is called “fattening”. The human brain and ears are not quick enough to hear two sounds at this delay time, humans perceives it as a “fatter” sound. The threshold of hearing one sound and two sounds are varied depending on the duration of the sound that is being delayed.

As well as reverbs, “fattening” is one of the most used effects in the studio, mostly because it doesn’t sound like an effect. “Fattening” is the primary effect used to make a sound stereo. When putting a dry sound in one of the speaker and put the delay less then 30ms in the other speaker, the sound “stretches” between the speakers. “Fattening” can be used to make a beautiful instrument more beautiful, and make a thin and irritating sound feel “fatter” and fuller and are the most efficient delay time of all the delay times to make this. It also makes the sound seem more present because it takes up more space when it is in stereo.

When using this short delay time phase cancellation will appear which is a big problem. The results of these cancellations are a loss of volume lost in mono. Bass frequencies will also be lost, clarity and precision of the perceived image of the sound will be lost. But if a sound is recorded with two microphones and phase cancellation is present then a delay set at 1ms can make the sound in phase.

5. Post processing of vocals: Interview analysis

The intent of the following analysis is to get a deeper understanding of vocal mixing and discover clues into how mixing engineers post process vocals and generally mix low-dynamic music. With this information extracted, the mix of music arrangement and post processes the vocals stimuli are determined.

To conduct the comparison of the post-processed vocal with the non post-processed vocals and get a understanding in how engineers uses effects on vocals when mixing, an analysis of the interviews made by B.Owsinski in his book “The mixing engineer’s handbook” [5] with leading mixing engineers and also for getting information on how to mix and post process the stimuli for the evaluation of the environments.

Eleven interviews with engineers were selected for the analysis based on the criteria that the engineers worked previously with rock and pop artists. The following engineers were selected and their credits according to the preface of the “The Mixing Engineer’s Handbook” are:

Joe Chiccarelli: Engineer and producer, worked with artists as Frank Zappa, Hole, U2, Beck, Tori Amos, Bob Seger

Lee DeCarlo: Engineer, worked with artists as Aerosmith, John Lennon, Rancid, Zak Wylde

Benny Faccone: engineer, worked with artists as, Mana (Latin rock band), Luis Miguel, Toni Braxton, Sting

Jerry Finn: Engineer, worked with artists as, Green Day, Rancid, Goo Goo Dolls, Beck


Kevin Killen: Engineer, worked with artists as, Stevie Nicks, Bryan Ferry, Pat Smith, U2, Elvis Costello
Allen Sides; Engineer, worked with artists as, Tevin Campell, Little Richard, Aretha Franklin

Don Smith; Engineer, worked with artists as, The Rolling stones, Tom Petty, U2, Stevie Nicks, Bob Dylan, Talking Heads, Iggy Pop, Keith Richards, The Eurhythmics, Cracker

Guy Snider; Engineer, worked with artists as, Ike and Tina Turner, Chuck Berry, Nine inch nails, Faith No More.

Ed Stasium; Engineer and producer, worked with artists as, Smithereens, Living Color, Mick Jagger, Soul Asylum, Ramones, Möterhead

Bruce Swedien; Engineer, worked with artists as, Jackie Wilson, Michel Jackson, Mick Jagger, Paul McCartney, Nat ”king” Cole and many more

5.1 Analysis of the interview
The goal of this stage is to search for similarities in their mixing process to see if a majority of the engineers were using a similar approaches in the use of effects on vocals, and what amount is used and why?
To draw clear conclusions and get manageable data of this analysis, the question was broken down in to categories of effects.
These questions/categories were chosen to get detailed, comparable quantity data (yes/no, and the overall usage of the effect much/less etc.) and qualitative data (why an effect was used/ to attain what sort of sound) for the analysis.

- **Does the engineer work with Equalizers?**
  
  Yes/no; under recording/mixing?; Why?; How much?; To attain what sort of sound?; On what?; What kind?

- **Does the engineer work with Compression?**
  
  Yes/no; under recording/mixing?; Why?; How much?; To attain what sort of sound?; On what?; What kind?

- **Does the engineer work with Delay?**
  
  Yes/no; under recording/mixing?; Why?; How much?; To attain what sort of sound?; On what?; What kind?

- **Does the engineer work with Reverb?**
  
  Yes/no; under recording/mixing?; Why?; How much?; To attain what sort of sound?; On what?; What kind?

- **Does the engineer’s work with Natural ambience:**
  
  Yes/no; under recording/mixing?; Why?; How much?; To attain what sort of sound?; On what?; What kind?

These questions are stated in a table along with the engineers on the Y axle and questions on the X axle to get a more complete overview of their opinions, see the table 12 in the appendix.
After evaluating the interviews with the help of the questions, the data points of each question were inserted in the table beside the corresponding engineer. To get a clear overview, the data
points from the tables are shown in graphs which are shown below, one question/category at a time.

**Question 1;** Does the engineer work with equalizers?  
Yes/no; under recording/mixing?; Why?; How much?; To attain what sort of sound?; On what?;  
What kind?

The usage of equalizer by the engineers;  
Eleven of eleven engineers are using equalizers.

![Figure 2.1](image-url)  
*Figure 2.1; The usage of Equalizer when Recording/Mixing or both.*
Figure 2.2: Illustrates why the engineers are using Equalizers.

Figure 2.3: The overall usage of Equalizers by the engineers when mixing.
To attain what sort of sound?

- Frequency balance
- Getting the feel of the instrument
- Making the instrument stand out
- Making the sound pleasurable
- Feel better

**Figure 2.4: The usage of Equalizers too attain a certain sound.**

The usage of Equalizer on different parts in the mix;
Eleven of the eleven data points from the analysis shows that engineers used equalizers on instruments and voice under the mixing process.

What kind of Equalizer is used?
Two out of two data points from the analysis showed that both engineers used frequency starting points when using equalizers.

**Question 2:** Does the engineer work with compression?
Yes/no; under recording/mixing?; Why?; How much?; To attain what sort of sound?; On what?; What kind?

**Figure 3.1: The usage of Compressor by the engineers.**

The usage of Compressor when Recording/Mixing or both;
Nine out of nine drawn answers from the engineers showed that the engineers use compressors when mixing, not under recording or both.

**Figure 3.3:** Illustrate why the engineers are using Compression, and shows that are no clear way of using the compressor, the data points are scattered.

**Figure 3.4:** The overall usage of Compression by the engineers when mixing.
To attain what sort of sound

- Breathing right
- Adding character to the sound or instrument
- Making sounds push through the track

Figure 3.5; The usage of Compression to attain a certain sound.

On what?

Answering engineers

Serie 1

Everything: 3
Drums: 1

Figure 3.6; The usage of Compression on different parts in the mix.
Question 3; Does the engineer work with delay?
Yes/no; under recording/mixing?; Why?; How much?; To attain what sort of sound?; On what?; What kind?

The usage of Delay by the engineers;
Nine out of nine data points (yes/no data points) drawn from the engineers in the analysis shows that all are using delay in some stage of recording, mixing or both.

The usage of Delay when Recording/Mixing or both;
Nine out of nine data points drawn from the analysis showed that the engineers are using delay only in the mixing stage.

Figure 3.7; Illustrates what kind of Compression is used.

Figure 4.3; Illustrates why the engineers are using Delay.
Figure 4.4; The overall usage of Delay by the engineers when mixing.

Figure 4.5; The usage of Delay too attain a certain sound.
The usage of Delay on different parts in the mix;
Six out of six drawn data points from the analysis shows that delay are used by the engineers primarily on vocals.

![Bar chart showing types of delays used](image)

**Figure 4.7; Illustrates what kind of Delays are used.**

**Question 4;** Does the engineer work with reverb?
Yes/no; under recording/mixing?; Why?; How much?; To attain what sort of sound?; On what?; What kind?

The usage of Reverb by the engineers;
Nine out of nine drawn data points from the analysis shows that all nine engineers are using reverb of some sort.

The usage of Reverb when Recording/Mixing or both:
Nine out of nine drawn data points show also that all engineers are using reverb in the mixing stage, not when recording or both.

Why the engineers are using reverb;
Only two data points could be drawn from this question in the analysis and both engineer’s answers has two different reasons of why they use reverb. To give the right character to the vocals which compliments it better and one of them are going throw a phase with less reverb.
The usages of reverb to attain a certain sound;
The pie chart shows that four data points were drawn from the analysis and that all have different reasons of using reverb, but three are using reverb too add some characteristic to the sound.

The usage of Reverb on different parts in the mix;
Three out of three drawn data points from the analysis shows that the engineers are using reverb on vocals when implementing it to the mix.
Figure 5.7; Illustrates what kind of Reverb is used.

**Question 5:** Does the engineer’s work with natural ambience?
Yes/no; under recording/mixing?; Why?; How much?; To attain what sort of sound?; On what?; what kind?

The usage of natural ambience by the engineers;
Two out of two drawn data points from the analysis shows that two engineers are using natural ambience.

The usage of Natural ambience when Recording/Mixing or both;
One engineer are using the natural ambience when mixing and the second engineer are using it in both the recording and mixing stage.

Why the engineers are using Natural ambience;
One data point could be drawn from the analysis. The reason why the engineer are using natural ambiance is to add reflections to the sound.

The overall usage of natural ambience by the engineers when mixing;
Out of the two data points that were drawn showed that both engineers are using natural ambience much when mixing and recoding.

The usages of Natural ambience too attain a certain sound;
Both engineers are using natural ambience to attain a natural sound to the mix.

The usage of Natural ambience on different parts in the mix;
Both engineers are using natural ambience on instruments only.

What kind of Natural ambience is used?;
The two engineers are using the natural ambience differently from each other, one is using first reflections and the other engineer is using the ambiance that follows with the recording.
5.2 Conclusions of the interview analysis

This section summarises the data from the graphs in the previous section and is followed by interesting findings from the interview analysis and a discussion section where each effect is discussed.

**Does the engineer work with Equalizers?**

- Eleven of the engineers use equalizers.

- Eight engineers use equalizers under the mixing process, one used it under the recording session and two used equalizers under both the recording- and mixing-session.

- The question “why the engineers use equalizers” illustrated that two of the engineers used to make the sound, sound natural. Two engineers used it to separate sounds, one used to make the sound feel better, one used the equalizer as a last resort, one used it to smooth the sound out, one used it to get rid of leakage between sounds and one did rather move the microphone then utilize equalizer.

- “The usage of equalizer much/less depending on the project” question illustrated that the overall usage of equalizer is; five uses equalizer much, five uses equalizers less and one uses equalizer much and less depending on the project.

- “The usage of equalizers is too attain a certain sound”, one used equalizers to get frequency balance, one used it to get the feel of the instruments, one used equalizer to make the instrument stand out, one used it to make the sound pleasurable and one used it make the sound feel better.

- Eleven engineers used equalizers on all instruments and voices.

- The kind of equalizer approach that is used by the engineers is using it with frequency starting points.

Eleven of the engineers use equalizers and a greater part of these engineers used it when mixing, the other part used it when recording and mixing. This pattern indicates that the use of equalizers is vast in some stage of the recording to mixing session. All the engineers used equalizer too process the sound in some way, ranging from getting rid of leakage between from other sound source, making the sound “feel” better and separate sounds. The sounds that the equalizers are used on are also vast, ranging from voices to instruments. Two engineers are using frequency starting points when they are using equalizers.

**Does the engineer work with Compression?**

- Eight engineers use compression, one does not use compression.

- Eight engineers use compression under the mixing process, All use it during the recording process or for both recording and mixing.
One engineer used compression too get the drum sound out of the way of the vocals, one used compression to get the sound controlled, smooth and consistent. One engineers used it because of demand from artist and producer, one used it the modify the sound, one used it to a lesser extent because of compression in later steps after mixing, one used it the add sustain without killing the attack, one used compression to add more ambience too the initial sound and one used it to add tones to the instrument.

“The usage of compressor much/less depending on the project” question illustrated that the overall usage of compression is; six engineers used compression much, two used it to a lesser extent and none used it much and less depending on the project.

The usage of compression is to attain a certain sound shows that one engineers uses it to make the sound breathe right, two uses it to add character too the sound or instrument and one uses it to making the sounds push through the track.

Three engineers used compression on everything in the mix and one used it on drums only.

The kind of compression approach that is used are; Compressor on a bus used by five engineers, Limiters and gates used by one engineer, compression on everything individual used by one engineer, and using the compression in a non consistent way was used by one engineer.

Eight out of nine drawn data points indicates that also here a vast majority of engineers are using compression especially under the mixing stage. However there are no consistent way of using the compressor, every one of the engineers are using it differently from each other. The amount of usage are also scattered but more then half of the engineer are using compression much. Why they are using compression question, indicates that the compressors major function in mixing are too add characteristic and make the sound too stand out in the mixing arrangement. Three engineers used it on everything in the mix. One used it on drums only and examining how the engineers use the compressor shows that there are also here no consistent way of implementing the compressor to the sound except from using compressor on a bus to add to the initial sound.

**Does the engineer work with Delays?**

- Nine engineers use delays.

- Nine engineers use delays under the mixing process and zero engineers uses it under the recording process or both.

- Two engineers use delays to make the vocal stand out from the arrangement, one uses it as pre delay to a chamber and one uses it to make the vocal “fatter”.

- “The usage of delay much/less depending on the project” question illustrated that the overall usage of delays are the following; eight engineers uses delays much, one engineer uses it less and zero engineers uses it much and less depending on the project.

- The usage of delays to attain a certain sound shows that two uses delay to making the vocal stand out from the arrangement and one uses it to make the vocal “fatter”.

- Six engineers use delays on vocals.
• The kind of delay approach that is used are; a delayed sent to another effect, used by two engineers; delayed sent to another effect in the tempo of the song, used by one engineer; tape machine delay, used by two engineers; Tape slap in the tempo of the song, used by two engineers; short delay, used by one engineer; delays in the tempo of the song, used by one engineer.

The delay are used by two engineers too attain the sound that makes the sound stand out of the arrangement and one other engineer wants to create the vocal sound to be perceived as “fatter”.

A majority of the engineers are using delay on the vocals and a variety of different implementations are used by the engineers when adding delay to the sound for creating the desired effect. It is clearly stated that the use of delay it too add/create a sound that stands out of the arrangement.

**Does the engineer work with Reverbs?**

• Nine engineers use reverb.

• Nine engineers use reverb under the mixing process and zero uses it under the recording process or both.

• One engineer use reverb too give the right character to the vocals too compliments the vocals better, and one engineer is going throw a phase with the use of less reverb.

• “The usage of reverb much/less depending on the project” question illustrated that the overall usage of reverbs is, three engineers uses reverb much, two uses it too a lesser extent and two are using much and less depending on the project.

• The usage of reverbs too attain a certain sound shows that one engineer is using it too add personality to the sound, one is trying the get the sound dry, one uses it when he wants it too sound lush and one is trying too get a room sound with reverb.

• Three engineers are using reverb on vocals.

• The kind of reverb approach that is used are, Plate reverb used by five engineers, reverb chamber used by one engineer, live chamber used by one engineer, echo chamber used by one engineer and non line reverb used by one engineer.

These results show that there is not consistent way of using reverb however the engineers can be divided in to categories, one category are the one that want too influence the sound and make is too sound special, lush and personal, and the other category which want too imitate the room feeling of where it was recorded. The results shows also that plate reverb are one of the most preferred reverb effects used by the engineers.

**Does the engineer work with Natural ambience?**

• Two engineers use natural ambience.

• One engineer is using it under the mixing process and one is using it under both the mixing- and recording-process.

• One engineer is using natural ambience too add first reflections to the sound.
• “The usage of natural ambience much/less depending on the project” question illustrated that the overall usage of natural ambiance is, two engineers that uses natural ambiance much.

• The usage of natural ambiance too attain a certain sound shows that one engineer is using it too create a natural sound.

• Two engineers are using natural ambiance on instrument.

• The kind of natural ambiance approach that is used are, first reflections used by one engineer, and using the ambiance that follows with the recording used by one engineer.

This data shows that there are few engineers that are using this practice, and it is used both under the mixing stage and the recording stage. It is for adding first reflections too the sound, however is it only used on the instruments. This could be because of lack of knowledge from the engineers that do not implement it or lack of experience with mixing with out of ambience effects.

5.3 Interesting findings of the interviews

Does the engineer work with equalizers?
• The majority of the engineers used equalisers when mixing.

• Two principles that are used by more than one engineer: using the equalizer too separate sound and to use it too make it sound natural.

• The engineers used equalizers on voice and instruments.

• Two engineer’s hade starting points when using the equalizer on instruments.

• For the question of what sort of sound is too attain, the data was scattered, no similarity was found.

The majority of the engineers used equalizers when mixing was not surprising due to the common practice of using equalizers too affect the sound in some way by boosting or cutting frequencies. The philosophies that are used more the one engineer is also not so surprising, because also common practice of using equalizers.

Some engineers used frequency starting points could be because of experience and knowledge of what type of instrument or voice type correlates generally and lies in special frequency bands. Too attain what sort of sound question, the data was too scattered this was expected due to the different way the engineers use equalizers and depending on what sort of sound ideal the engineers wants too achieve with that particular sound.

Does the engineer work with Compression?
• Nine of the engineer’s use dynamic post processing, eight of them use compression, and the one engineer uses limiters and gates when mixing.

• A majority of the engineers is using compression when mixing and none of them when recording
• The data indicates that a vast majority uses it much when mixing.

• Three of them used individual compressors on every sound source when mixing.

• One major approach was the use of compression over a bus to add more character to the sound. The other approaches were scattered, by compressing each instrument and vocal individually, or using it by trial and error.

It was expected that all engineers used some amount of compression due to the designed purpose of the compressor stated in the background section, however that one engineer used limiter and gates but no compressor is surprising, these equipments are essentially, including the compressor, a part of the dynamic sound processing equipment family. A majority used to compression only in the mixing stage, this could be to not add anything that changes the dynamics of the sound before it is recorded. A third of the engineers used compression on everything, this result was not expected, this gives an hint that compression are effective in low-dynamic music for controlling sounds. The major approach on how to implement the compressor was to have a compressor over a bus which is the added to the initial sound, this was not expected, but it illustrated the compressor can also be used as effect instead as a dynamic tool.

Does the engineer work with Delays?

• Nine engineers uses delays when mixing, eight of them uses delays much.

• The dominant use of delay is on the vocals to either add character to the sound which is the major use, or to “fattening” the sound. To accomplish this, the delay should be used in the tempo of the song.

That the engineers are using delay to a large degree was not expected due to the expected dry nature of low-dynamic music. The use is mostly on vocal which indicates that vocal sound in low-dynamic music should not be dry. The delay time should be in the tempo of the song were not expected, due to the dry fullness expected from the vocals.

Does the engineer work with Reverbs?

• The data drawn from the interviews shows that nine of the engineers are using reverbs under the mixing process.

• The amount of usage is scattered, however the large amount of usage is in favour versus the less-, much and less depending on the project-usage.

• The sound source that the reverbs were used on was vocals.

• Why the engineers were using reverbs indicates that it gives the right character to the vocals which compliments it better and the use of less reverb can be because of a personal phase from the engineer.

• To attain a certain sound the data points was scattered, no one hade the same qualifications on the same sound.

• Plate reverb is the most used reverb among the engineers.
That the reverb were only used in the mixing process was expected can be because of the favour of unprocessed material before mixing, a majority of the engineers used reverbs mostly on vocals, this was not expected due to the dry fullness of the low-dynamic music.

“Why the engineers used reverb on vocal” indicated that it was applied to give the vocals the right character, this was expected, due to the subjective liking of each one the engineers.

On how to attain a certain sound with the reverb the data was scattered this was also expected because the subjective liking of the engineers. One data point that was not expected, the plate reverb were the most used reverb amongst the engineers, this can be because of size of the plate unit which are smaller then an echo chamber or due to the liking of the engineers.

**Does the engineer work with Natural ambience?**

- Two of the engineers started working on the recorded ambient (or ambient microphones) which follows every recorded instrument, before they reconsidered the use of reverb effects.

- They used natural ambience much, but the use was mainly on instruments.

- They used it to add first reflection of the sound, and used the ambience that follows with the recording.

That two of the engineers started working on the recorded ambient (or ambient microphones) which follows every recorded instrument, before they reconsidered the use of reverb effects was not expected, this gives the indication that this way of working are a part of their personal working style. The use of natural ambience on only instruments was expected due to the lack of information on the use of natural ambiance on vocals in musical recordings. The use to add first reflection of the sound, and the use of the ambience that follows with the recording were also unexpected due to the lack of information on how to use natural ambiance in musical recordings.

**5.4 Discussion on the interview analysis**

One thing that was detected under the analysis, that the data didn’t show was; the amount of effects added was determined according to the musical style. If there is a pop song, a bigger amount of reverb on vocals might be used and when dealing with rock song a lesser amount of reverb might be used on the vocals and that the focus point of the engineers was on the vocals and it should correspond to the music, this might be an indicator that in dynamic music the listener of the song prefers more reverb, then the listener would prefer in low-dynamic music.

Since the approach on mixing was so varied from engineer to engineer, here is an observation on the parameters affecting the choices in each case, a majority of them had an initial idea that the rhythm section (bass and drums) and vocals should be treated first, then add the other instruments in the arrangement and mix them into the song.

These statements below illustrates that the vocal is important in the mixing phase from the artist point of view, to the get the lyrics across, and also for the frequency placement of the other instruments so that they don’t collide with each other. Another part that the engineers were focussed on, was the part in the song that makes it “tick” however this data was to scattered between the engineers and were to subjective, and this work is based on vocal sound, further analysis of this part was withdrawn.

---

2 A dominant instrument or melody that is in the song.
Guy Snider; "I throw up all the faders and listen to what the song I about, and then I start with drum sounds. Then I throw the vocal in. The vocal’s always sort of the barometer. Plus I use it also for frequency placement. A lot of the time the snare is fighting for the same place as the vocal, so I always make sure that the snare sound really good…….punched back into the mix"(Owsinski 1999, s 186) [5].

Don Smith; "Most of the time just start with the drums and bass, then everything else. Then there were some records that I started with lead vocal…..With somebody like Tom Petty, his vocals is so important in the mix that you have to start with the vocal"(Owsinski 1999, s 181) [5].

Jerry Finn; "Lately, I’ve tried to put the vocal in early in order to create the mix more around that. In a lot of the Punk Rock stuff you get the track slamming and then you just sort of drop the vocals on top. But with poppier stuff I’ve found that that approach doesn’t work as well because the vocal really need to sell the song"(Owsinski 1999, s 112) [5].

The use of Equalizers
None of the engineers wanted the audio to feel “unnatural” when they used equalizing. The term natural and smooth were frequently used to describe the sound that they wanted when they used equalizers, which indicates that the use of equalizers can be used a lot but not to a degree of degradation of the sound, this is also indicated in the data which can lead to the theory that the equalizer is use to enhance the sound.

The use of Compressor
The use of the compressor can be seen as an effect to control the sound and make it stand out from the rest of the sounds sources. When interpreting Benny Faccone’s statement.

Benny Faccone; "Limit the heck out of everything [laughs] , I like to compress everything just to keep it smooth and controlled, not to get rid if the dynamics, But I don’t like a compressor across the stereo buss because then it’s sounds like it’s not breathing right to me. Even for hard rock, I don’t like to do that. It’s easier to do it individually……on the bass I hit that a little harder, just to push it up front a little more, everything else for control more than sticking it right up in your face kind of a thing.”(Owsinski 1999,s 109) [5].

This gives a hint that compression can be used to lift an instrument or vocal to stand out from the rest of the musical material, and it also gives a second approach on how to use the compressor in stead of the majority buss compressor approach.
All the engineers speaks positively to compression in the interviews, as a way to get good control and even to an extent that Jerry Finn said "I think that the sound of modern records today is compression"(Owsinski 1999,s 114) [5]. This is just one statement of a person’s opinion but it illustrates a seemingly general attitude toward compression as seen in all the engineers.

The use of Delay
This statement from Don Smith is about delays on drums but it shows that delays are being used only for giving an notion that there are a delay effect on the sound, so that the delay are merely audible by the human ear, as the data indicates. The data indicates also that delay is used to make the sound stand out more and draw the listener’s attention. This can be why it is frequently used on vocals and why the listener is drawn to the vocals in a song. Almost every engineer used delay in some way and had a different approach which indicates that there is no right or wrong use of delay, if it is added in tempo of the song or with a very short delay time.
Don Smith: "Sometimes on the drums I’ll use delays very subtly. If you can hear them, then they’re to loud but if you turn them off, you defiantly know that they are gone, it adds a natural slap like in a room, so to speak, that maybe you wont hear but you feel" (Owsinski 1999, s 182) [5].

The use of Reverb
The data points on why the engineers used reverb was because reverb gives the right character to the vocals which compliments it better and the use of less reverb can be because of a personal phase, these data points are vague because there were only two opinions that could be drawn from the interviews and has no support from the rest of the engineers.

The data from “to attain what sort of sound” indicates that reverbs are used as enhancers to the sound source, which indicates that reverb is also used on vocals with delay to enhance the vocals so that the listener are drawn to the vocals in the song. The type of reverb that is primary used on vocals is plate reverb, however, three of the engineers used some sort of chamber to create reverberation, however the difference in usage and the choice between the plate reverb and the other reverbs was selected based on the premises.

The use of Natural ambience
Only two engineers worked with natural ambience and the use was primarily on instruments, this indicates that using the real ambiance is not a well established work method amongst the mixing engineers however the engineers worked much with the natural ambience and can be perceived as a method hard to master or as a forgotten art in the mixing industry. When reading the data it is clear that it is not a common practice.

6. Mixing Constraints for the vocal stimuli’s and musical arrangement

In this section mixing constraints for the mix of the stimuli are drawn based on the findings form the interview analysis. The mix of the instruments is identical for all stimuli’s and reference stimuli, the only thing that are variable in the mix are the vocals, because the vocal sounds are the focus point in this project. A detailed list over the recording and mixing equipment see table 13 in the appendix. The mixing of the stimuli was made by the researcher/author with knowledge attained from the interview analysis.

Stimuli arrangement details
The song is a low-dynamic rock/pop song with the emphasis on rock with a tempo of 140 bpm. The instruments in the arrangement are: Drums, bass, guitar, piano and tambourine.

Frequencies roughly range of the individual instruments before mixing:

Guitar: 100-7000 Hz
Bass: 20-200 Hz, 5 – 7 kHz
Piano: 100-8000 Hz
Tambourine: 100-20 000 Hz
Drums:
  Kick drum: 20-3000 Hz
  Snare drum: 200-10 000 Hz
  Toms: 40-5000 Hz
  Cymbals and high hat 40-20 000 Hz
**Vocal stimuli details**
Room dimensions can be seen under the topic: Methodology. Decay time and first reflections/pre delay times can be seen in table 8-11 in the appendix.

Vocal recorded in Song booth (K9).
Vocal recorded in an anechoic chamber.
Vocal recorded in a small and narrow entrance hall to an apartment.
Vocal recorded in a church.
Vocal recorded in a song booth (K9) with post processed ambience effects.

**Vocal mix of the stimuli’s**
The vocals will have constant sound amplitude on the fader in all four stimuli’s, this level will be set by the mixed post processed vocals, the same amplitude in all stimuli’s. The vocals are not panned or modified in any way, only for making the evaluation of the stimuli’s clearer.

Placement of the vocals: Centre

**Vocal mix of the post processed stimuli**
The post processed vocal will have a send with a plug-in\(^3\) plate reverb with a pre delay time about 30ms to lift the vocal and the plate reverb with a short reverb time so it doesn’t take up a lot of space in the sound image. The plate reverb is added subtly to the initial vocal sound. Table 1 illustrates the setting on the post process plate reverb. All the fader levels for instruments and vocals are set by the researcher/mixer to achieve the low-dynamic aspect of the mix. The fader levels used in the mix can be seen in table 2.

Placement of the vocals: Centre

---

**Table 1: D-verb settings**

<table>
<thead>
<tr>
<th>D-Verb</th>
<th>Settings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input</td>
<td>0dB</td>
</tr>
<tr>
<td>Mix</td>
<td>100%</td>
</tr>
<tr>
<td>Diffusion</td>
<td>40%</td>
</tr>
<tr>
<td>Decay</td>
<td>401ms</td>
</tr>
<tr>
<td>Pre Delay</td>
<td>30ms</td>
</tr>
<tr>
<td>High Frequency cut</td>
<td>OFF</td>
</tr>
<tr>
<td>Low Pass Filter</td>
<td>OFF</td>
</tr>
<tr>
<td>Algorithm</td>
<td>Plate-Large</td>
</tr>
</tbody>
</table>

**Mix of the musical arrangement for the stimuli**
To avoid extensive collisions\(^4\) between the instruments frequency wise, equalizers are used on Piano, Bass and overhead tracks. For getting control over the instruments and making them not to collide with the vocals, compression is used on the snare drum, bass drum, toms, over head microphones and tambourine. The volume of the instruments is not to mask the vocals, merely to fighting with the vocals to achieve the low-dynamic aspect of the arrangement.

---

\(^3\)Digital version of a plate reverb, which imitates a real plate reverb functions.

\(^4\) Collisions of frequencies are unavoidable however minimising the effect of collisions can give a desired result.
Placement of the instruments:
The drums and bass are placed in the centre of the sound image, the guitar are panned to right and
the piano to the left. The tambourine is also placed towards the right. This is for making room for
the vocals.

Figure 6; illustrates the placement of the instruments and vocals in the sound image between the
speakers horizontally, perceived from the listener’s point of view.

Table 2; Shows the fader levels, type of plug-Ins and microphones on each track of the mix.

<table>
<thead>
<tr>
<th>Track</th>
<th>Fader levels</th>
<th>Inserts</th>
<th>Microphones</th>
</tr>
</thead>
<tbody>
<tr>
<td>BassDrum1</td>
<td>- 4dB</td>
<td>Compressor</td>
<td>Sennheiser MD 421</td>
</tr>
<tr>
<td>BassDrum2</td>
<td>- 8.6dB</td>
<td>Compressor</td>
<td>Shure D112</td>
</tr>
<tr>
<td>Snare Drum</td>
<td>-1.0dB</td>
<td>Compressor</td>
<td>Shure SM 57</td>
</tr>
<tr>
<td>Toms</td>
<td>-1.2dB</td>
<td>Compressor</td>
<td>Sennheiser MD 421</td>
</tr>
<tr>
<td>Overheads</td>
<td>-1.0dB</td>
<td>1 band Eq, Compressor</td>
<td>L/R 2° AKG c414</td>
</tr>
<tr>
<td>(Stereo)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drum Master</td>
<td>-3.2dB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bass Amp</td>
<td>-2.6dB</td>
<td></td>
<td>Shure SM 57</td>
</tr>
<tr>
<td>Bass D.I.</td>
<td>+ 2.2dB</td>
<td></td>
<td>BSS Audio AR-133 active D.I. box</td>
</tr>
<tr>
<td>Bass Master</td>
<td>-4.7dB</td>
<td>4 band Eq, Compressor</td>
<td></td>
</tr>
<tr>
<td>level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tambourine</td>
<td>-17.0dB</td>
<td>4 band Eq, Compressor</td>
<td>AKG c414</td>
</tr>
<tr>
<td>Guitar</td>
<td>-5.4dB</td>
<td>1 band Eq</td>
<td>Shure SM 57</td>
</tr>
<tr>
<td>Piano (Stereo)</td>
<td>-8.0dB</td>
<td>4 band Eq</td>
<td>Left, AKG c414 bottom of piano</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Right, Neumann KM 184 strings of piano</td>
</tr>
<tr>
<td>Anechoic</td>
<td>-1.9dB</td>
<td></td>
<td>AKG c414</td>
</tr>
<tr>
<td>chamber Vocal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entrance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hallway Vocal</td>
<td>-1.9dB</td>
<td></td>
<td>AKG c414</td>
</tr>
<tr>
<td>K9 Vocal</td>
<td>-1.9dB</td>
<td></td>
<td>AKG c414</td>
</tr>
<tr>
<td>Church Vocal</td>
<td>-1.9dB</td>
<td></td>
<td>AKG c414</td>
</tr>
<tr>
<td>Aux Reverb</td>
<td>-13.0dB</td>
<td>D-verb</td>
<td></td>
</tr>
</tbody>
</table>
7. Listening Test

The purpose of the listening test is to find out if there is an advantage in recording the vocals in a real environment with natural reverberation instead of post processing the vocal with ambiance effects. The five subjects who participated in the test have an education in audio engineering, and do not suffer from any form of hearing disorder. This data was retrieved on the form before the evaluations of the stimuli’s were started. The use of educated subjects was to conduct a test in which the subjects could give a precise explanation in words of what they perceives from each environment, then the explanations that laymen would have given. The subjects did the test anonymously to maintain the experiment’s integrity.

The types of data that examined in this listening test are;

- How are each vocal perceived in each one of the mix including the post processed vocal (reference)?
  - Intelligibility of the lyrics.
  - Volume of the vocals, which is perceived loud or low versus the music arrangement.
  - The amount of reverberation sound of the vocals versus the direct sound of the vocals i.e. which stimuli has what amount of reverberation proportion.

This is for getting an understanding of how each vocal stimulus are perceived in the arrangement by the subjects, and to see if there are variations in the perception between the stimuli’s from the subject’s/listeners point of view.

- How the subjects perceives the size of the environments where the vocals are recorded and if there are any positive or negative feature with these five environments.

- Which one of the 5 vocal environment is mostly preferred in the mix and why?

These questions have a qualitative approach to get a more subjective perspective on how each stimulus are perceived from the subjects/listeners point of view.

The subjects are informed that he/she are conducting a listening test were they are evaluating five different vocals sound in low-dynamic music and how they perceive these vocal environments. They are given the instruction to read the form first, before they start the test. This emphasizes that it is the vocal environments that are at the centre of the evaluation, and also this guides thire attention toward the vocal sound and not the other parts of the mix.

Here are the questions shown which are stated on the form, the instructions for the subjects and the entire form can be seen in the appendix.

- How are each lead vocal perceived in each one of the mix?
  - Intelligibility (tydlighet) of the lyrics
  - Volume of the vocals, which is perceived loud or low versus the music arrangement.
  - The amount of reverberation sound of the vocals versus the direct sound of the vocals i.e. which stimuli has what amount of reverberation ratio

- Rank each one of the stimuli’s in order of perceived room size

- Did the perceived room size affect the vocals positive or negative in the mix?
Stimuli A: Positive  Negative  Why?
Stimuli B: Positive  Negative  Why?
Stimuli C: Positive  Negative  Why?
Stimuli D: Positive  Negative  Why?
Stimuli E: Positive  Negative  Why?

- Rank the 5 lead vocal environments from mostly preferred to the least preferred by you in the mix?
- Why is it the most preferred?
- Why is it the least preferred?

7.1 Listening test implementations

In the listening test five vocal stimuli are played using an application coded in the Java language which gives the subject full control over the playback of the stimuli, on a computer. The subject conducting the test can play each stimulus as many times as he/she prefer when answering the questions on the form. The test is preformed in an anechoic lab in The Technical University of Luleå. This laboratory has no background noise, no reflections and no decay time. The subjects are given closed studio headphones, for eliminating the sound of the computer under the listening test. The headphones have a fixed volume setting through out the entire test. The volume is set to 0dB which is the maximum in the software mixer to the soundcard, this level are defined appropriate by the researcher, due to software mixers monitor level only goes to 0dB, which were a comfortable level, however, if one or more subjects desires a louder listening level then there were no further headroom on the monitor level. There were no complaints by the subjects on the listening level after the test.

To control over the order of the stimuli, in the listening test with a small number of participants the Latin square is used to arrange the numbers/letter in a matrix, so the number/letter occurs once in each row and column [6]. The participant will have the stimuli in different orders to minimize the comparison by the subjects after and before the test with other subjects and to avoid that the result is due to the order of the stimuli [6]. The order of the stimuli will be different in each test, however the subject will see the order on the screen as A B C D E, even if the actual order is of the stimuli’s are B A E C D.

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Order</th>
<th>Perceived order</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A B C D E</td>
<td>A B C D E</td>
</tr>
<tr>
<td>2</td>
<td>B A E C D</td>
<td>A B C D E</td>
</tr>
<tr>
<td>3</td>
<td>C E D A B</td>
<td>A B C D E</td>
</tr>
<tr>
<td>4</td>
<td>D C B E A</td>
<td>A B C D E</td>
</tr>
<tr>
<td>5</td>
<td>E D A B C</td>
<td>A B C D E</td>
</tr>
</tbody>
</table>

Figure 7: Stimuli order for each subject.

The equipment used when conducting the listening test are shown in table 3.
Table 3

<table>
<thead>
<tr>
<th>Listening test equipment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laptop</td>
</tr>
<tr>
<td>Headphones</td>
</tr>
<tr>
<td>Soundcard</td>
</tr>
<tr>
<td>Computer interface</td>
</tr>
</tbody>
</table>

7.2 Evaluation of the stimuli’s in the listening test

The first question of the listening tests is divided into three subcategories for getting manageable data on how each stimulus are perceived. The subcategories are graded by the subjects on ordinal scales, the ordinal scale is used for getting the magnitude of preferences from each one of the subjects on the stimuli’s parameters and there by establishing how each stimuli’s intelligibility, volume and reverberation are perceived [6]. The scale is ranging from 1-5, with an interval of 1. A one is perceived as the worst/lowest intelligibility, the lowest amount of perceived volume and lowest amount of reverberation time on the vocals, a five is perceived as best/highest intelligibility and the highest amount of perceived volume and reverberation time.

The second question is a close-ended question which restricts the subjects to answer on a certain set parameters/fixed scale points. It is a ranking question on an ordinal scale, the same as the first questions scale [6]. The subjects are evaluating how they perceive the size of the room that the vocals are recorded in. A five is perceived as the large sounding room and a one is perceived as a small sounding room.

The third question is an open-ended question. The subjects can answer the question freely in their own words [6]. The subjects are evaluating if there are any positive or negative with the perceived room size and are given a chance to explain in their own words/writing what the positive or negative qualities are in each one of the stimuli.

The forth question is constrained question which allows the subjects to answer on which one of the stimuli are most preferred and not preferred by the subjects. The scale is an interval scale from 0,5-10 with an interval of 0,5 this is for seeing the subject spread of their liking, if one environment is closer another environment [6]. The fifth question is an open-ended question on why he/she chooses that particular stimuli as most preferred and least preferred. The form can be seen in the appendix.

The time each subject had to complete the test was one hour, the average time the subjects used was approximately 20-35 minutes. The subjects were told to read the form before starting the test and understand each question and if they had any questions then the researcher would answer them before the test.

Due to some complications, with crackle in the sound and buttons stopped working on the interface when the listening test was conducted, a reboot instruction was written after the first subject. The subject were told to close the interface, restart it and select there user id number, ranging from 1-5 which was stated on the first page of the form, if the subject did not understand the instructions then the researcher would show step by step how to restart the interface. The researcher asked afterwards if the rebooting had an affect on there performance, then the subjects answered it was quick to restart the interface and said it was a minor problem. Due to the subject’s answers and the unlimited times the subjects could listen and evaluate each stimulus, the listening test was continued. The number of times the subjects hade to restart the interface in each test was 1-3 times for all subjects. One opinion which was stated by several subjects after the test was that the stimuli hade to long musical segments and no fast forward or rewind buttons which caused the subject to wait for the vocals to begin.
7.3 Expected data from the listening test

The post processed vocal has a pre delay time on 30ms and reverberation with a diffusion set on 40\% and a decay time of 401 ms on an aux channel which are added to the vocal track at a level of\(-15\)dB and the vocal track at \(-1,9\)dB. This amount of ambience can be perceived as the vocal has been recorded in a not so damp, medium sized room, due to the threshold of detection of reflections, which shows that the spaciousness of the sound becomes audible when the reflected sound level reaches about 15dB or less below the direct sound level [3].

The anechoic environment would be perceived to have no reverberation time or reflection/pre delay at all and should be rated as the one with least amount of reverberation due to the non exiting room ambiance in the vocal sound.

The recording from the entrance hall has short reflection/pre delay time of 5 ms at a \(+50^\circ/-50^\circ\) horizontal angel from the sound source, and a reflection/pre delay time of 4 ms at \(0^\circ\) vertical, with the sound source. The decay time is of approximately 0.4 seconds in the room. This environment can be perceived as unpleasant due to comb filter effects on the sound based on the small reflection times from the walls, floor and ceiling [3].

The recording which was recoded in a church has a large decay time of approximately 1 s and a reflection/pre delay time of 0,028 ms at an angel of \(0^\circ\) from the sound source. This vocal sound can be perceived as very reverberant and due to the long decay time and long pre delay time in the room, this can influence the vocal and make the vocal sound more spacious because of the audible reverberation perceived by the researcher, which indicates that the reflections are less the 15dB below then the direct sound in and are over the threshold of detection of the reflections [3].

The recording which was recorded in the song booth has a decay time of approximately 400ms and has a reflection/pre delay time of 5 ms at a \(0^\circ\) angle with the sound source. The vocal sound can be perceived as low reverberant due to the damping in the room which was perceived under the measurements of the decay time. It is also shown in the results of the measurements (see table 9 in the appendix) that the decay time is almost the same for all octave bands, except for minor verities. These measurements indicate that the reflections are in someway controlled.

7.4 Data from the listening test

In this section are the data presented from the listening test, and short overviews of what might have influenced the subjects to give this data, a more detailed discussion of what have influenced the subjects and what parameters the subjects used to evaluate the environments and comparisons between the questions are discussed and drawn under the next topic.

• How are each lead vocal perceived in each one of the mix?
  o Intelligibility (tydlighet) of the lyrics
  o Volume of the vocals, which is perceived loud or low versus the music arrangement.
  o The amount of reverberation sound of the vocals versus the direct sound of the vocals i.e. which stimuli has what amount of reverberation ratio

• Rank each one of the stimuli’s in order of perceived room size

• Did the perceived room size affect the vocals positive or negative in the mix?
• Rank the 5 lead vocal environments from mostly preferred to the least preferred by you in the mix?

• Why is it the most preferred?

• Why is it the least preferred?

**Question 1;** How are each lead vocal perceived in each one of the mix?

This question is divided in to subcategories for evaluation different parameters, Intelligibility (tydlighet) of the lyrics, Perceived volume and amount of Reverberation. Each bar illustrates the ranking of the stimuli. As stated previously a five corresponds to the most amount of intelligibility, perceived volume and reverberation and a one illustrates the least perceived amount.

**Intelligibility of the lyrics**

![Intelligibility Graph]

*Figure 8.1: Ranking of most-least Intelligibility on each one of the stimuli perceived by the subjects.*

- The church environment was rated by two subjects as the second least intelligible and three subjects rated the church on grades from 3-5.

- The song booth got the biggest rating as the second most intelligible environment by three subjects. One subject rated the song booth (K9) as the most intelligible and one subject rated it as the second least intelligible environment.
o The entrance hall was graded by four subjects as the least intelligible environment and one subject rated the entrance hall as the second least intelligible.

o The anechoic chamber was rated as the most intelligible by two subjects, one subject rated it the second most intelligible, one subject rated it with a grade of 3 and one subject rated it as the least intelligible environment.

o The song booth plus the reverb environment were rated by two subjects as the second least intelligible and by two other subjects a ranking of 3, and one subject graded it as the most intelligible environment.

Figure 8.2; Ranking of most-least perceived volume at each one of the stimuli by the subjects.

o The church had two subjects that ranked the environment as the one with the most perceived volume and two subjects ranked it the second most perceived volume and one subject ranked it a 3 on the scale.

o The song booth (K9) was rated by two subjects as the second least amount of volume, one rated is as the one which have the second most perceived volume, one subject rated it with the least amount of perceived volume and one rated it a 3 on the scale.

o The entrance hall was perceived by two subjects as the one with least amount of volume and two other subjects rated it the second least amount of volume. One graded it as the second one with most perceived volume.
- The anechoic chamber was rated the grade 3 by two subjects, one subject rated it the one with the most perceived volume and two subjects rated as the one with the least perceived volume.

- The song booth (K9) + the reverb environment was rated by two subjects as the environment with the most perceived volume, one subject rated it as the one with the second most perceived volume, one rated it as the one with the second least perceived volume, and one rated it as the one with the least perceived volume.

**Amount of reverberation**

![Bar chart showing the ranking of the amount of reverberation at each environment by the subjects.]

**Figure 8.3:** Ranking of most-least perceived amount of reverberation at each one of the stimuli by the subjects.

- The church was rated by two subjects as the second one with most amount of reverberation, two subjects rated it as a 3 on the scale, and one subject rated it as the second one with the least amount of reverberation.

- The song booth (K9) was rated by three subjects as the second with least amount of reverberation, one subject rated it as a 3 on the scale, and one subject rated it the second one with the most amount of reverberation, this indicates that the song booth (K9) has a low amount of reverberation, except from one subject how strayed by rating it as the second most reverberant environment.

- The entrance hall had four subjects which rated it as the one with most amount of reverberation, and one subject rated as the second one with the least of reverberation.
The anechoic chamber was rated by all subjects as the one with the least amount of reverberation.

The song booth (K9) + reverb was rated by two subjects as a 3 on the scale, two rated it as the second one with most amount of reverberation, and one subject rated it as the one with the most amount of reverberation.

**Question 2;** Rank each one of the stimuli in order of perceived room size.

Each bar of the graph illustrates the subjects ranking of each environment with grade one, two, three, four or five which corresponds to the perceived size of environment. Grade five is the largest perceived size and a one is the smallest perceived size of the environment.

![Figure 8.4; Ranking of the perceived room size at each one of the stimuli.](image)

- The church was rated by one subject as the one as the largest room, three subjects rated in as a third on largest/smallest room size, and one rated it as the second one with the smallest room size.

- The song booth (K9) was rated by three subjects as the second one with the largest room size. One subject rated it as the third on largest/smallest room size, and one subject rated it as the second one with the smallest room size.

- The entrance hall was rated by three subjects as the one with the largest room size and two subjects rated it as the room with the smallest room size.

- The anechoic chamber was rated by three subjects as the one with the smallest room size and two subjects rated it as the second one with the smallest room size.
The song booth (K9) + reverb was rated by two subjects at the second one with the largest room size, one subject rated it as the largest room, one subject rated it as third on the largest/smallest room size and one subject rated it as the second one with the smallest room size.

**Question 3;** Did the perceived room size affect the vocals positive or negative in the mix?

- Four subjects perceived the church room size to affect the vocals positive. One subject perceived the room size to affect the vocals negative.
- Two subjects perceived the song booth room size to affect the vocals positive and two subjects perceived the room size to affect the vocals negative.
- All subjects perceived the entrance halls room size to affect the vocals negatively.
- Four subjects perceived the anechoic chamber size to affect the vocals negatively.
- Three subjects perceived the song booth plus reverbs room size to affect the vocals positive and two subjects perceived the room size to affect the vocals negatively.

Table 4 illustrates how many subjects perceived negative/positive affects of the room size on the vocal sound. The zeros indicate that no subject did perceive anything positive or negative from that environment and the answer on the form were left blanked. In table 5 are what the positive and negative affect presented.

<table>
<thead>
<tr>
<th>Positive/negative effects on perceived room size</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Church</strong></td>
</tr>
<tr>
<td>Positive</td>
</tr>
<tr>
<td>Negative</td>
</tr>
</tbody>
</table>

**Table 5:** illustrated what negative/positive effects are perceived by the subjects from the room size.

<table>
<thead>
<tr>
<th>Subject one</th>
<th>Positive</th>
<th>Negative</th>
<th>Why?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Church A</td>
<td></td>
<td>X</td>
<td>The reflections feels separated from the direct sound. The Pre delay time is to long.</td>
</tr>
<tr>
<td>K9 B</td>
<td>X</td>
<td></td>
<td>Good response, pleasant but not to audible. Better Swing.</td>
</tr>
<tr>
<td>Entrance Hall C</td>
<td>X</td>
<td></td>
<td>The room sounds shut in, assertive, comb filter effects.</td>
</tr>
<tr>
<td>Anechoic chamber D</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K9+ Reverb E</td>
<td></td>
<td>X</td>
<td>Good sound but too big for the mix, to clear.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subject two</th>
<th>Positive</th>
<th>Negative</th>
<th>Why?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Church A</td>
<td>X</td>
<td></td>
<td>The vocal fitted in the song.</td>
</tr>
<tr>
<td>K9 B</td>
<td>X</td>
<td></td>
<td>The room size gave a pleasant space for the vocals.</td>
</tr>
<tr>
<td>Subject three</td>
<td>Entrance Hall C</td>
<td>The vocal feels to dry.</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------</td>
<td>-------------------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>The vocal doesn’t feel as a part in the song.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Anechoic chamber D</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td></td>
<td>K9+ Reverb E</td>
<td>The room doesn’t feel small/Narrow</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subject three</th>
<th>Church A</th>
<th>Natural.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K9 B</td>
<td>Natural, but to hard and dry.</td>
</tr>
<tr>
<td></td>
<td>Entrance Hall C</td>
<td>Too much echo.</td>
</tr>
<tr>
<td></td>
<td>Anechoic chamber D</td>
<td>To dry.</td>
</tr>
<tr>
<td></td>
<td>K9+ Reverb E</td>
<td>Pleasant amount of room for this type of vocal and music.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subject four</th>
<th>Church A</th>
<th>Adequate tuning of the amount of reverb (room size) and intelligibility.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K9 B</td>
<td>To dry.</td>
</tr>
<tr>
<td></td>
<td>Entrance Hall C</td>
<td>To dry sounds like a tin and do not fit.</td>
</tr>
<tr>
<td></td>
<td>Anechoic chamber D</td>
<td>Clear but too dry and doesn’t fit in the mix.</td>
</tr>
<tr>
<td></td>
<td>K9+ Reverb E</td>
<td>To big and do not fit in the mix.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subject five</th>
<th>Church A</th>
<th>Good sound image where the direct sound and reverb/room is balanced with each other. Cleartess on the vocal.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K9 B</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Entrance Hall C</td>
<td>Too much room, can be unclear and a confused sound image.</td>
</tr>
<tr>
<td></td>
<td>Anechoic chamber D</td>
<td>To dry, ends up to close in the mix, one can lose focus on the other instruments, hard to understand the vocal sometimes.</td>
</tr>
<tr>
<td></td>
<td>K9+ Reverb E</td>
<td>The vocal ends up farther back in the mix, good balance between the instruments, good amount of reverb, makes it balanced.</td>
</tr>
</tbody>
</table>
**Question 4:** Rank the 5 lead vocal environments from mostly preferred to the least preferred by you in the mix.

Each bar of the graph illustrates the subjects ranking of each environment with grades a half to ten which corresponds to the preferred environment. Grade ten is the most preferred environment and a half is the least preferred environment.

![Preferred environment diagram]

*Figure 8.5: Perceived environment by the listener ranked on a scale from 0.5-10 where one is the least preferred and a ten is the most preferred.*

- Three subjects rated the church the most preferable (10) on the scale, one subject rated it a 7, 5 on the scale, and one subject rated a 3, 5 on the scale.

- The song booth (K9) was rated the most preferred by one subject (10) on the scale, one subject rated it 7, 5 on the scale, one subject rated it a 6, 5 on the scale, one rated it a 5, 5 on the scale and one subject rated it a 4, 5 on the scale.

- The entrance hall was rated by one subject a 2, 5 on the scale, one rated it and 1, 5 on the scale and three subjects rated it the least preferred environments (0, 5) on the scale.

- The anechoic environment was ranked by one subject a 6, 5 on the scale, one subject rated it a 5 on the scale, one subject rated it 3, 5 on the scale and two subjects rated it as the least preferred environment (0, 5) on the scale.

- The song booth (K9) + reverb was rated by one subject the most preferred environment, one subject rated it a 9 on the scale, one subject rated it a 8, 5 on the scale, one rated it as a 6, 5 on the scale and one subject rated it an 1,5 on the scale.
**Question 5; Why is it most preferred, Least preferred**

The answers can be seen in the table 6, were all the data are stated.

Table 6: $A =$ *Church*, $B =$ *Song Booth (K9)*, $C =$ *Entrance Hall*, $D =$ *Anechoic chamber* and $E =$ *K9+Reverb*.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Most Preferred</th>
<th>Why most preferred?</th>
<th>Least Preferred</th>
<th>Why least preferred?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject 1</td>
<td>B</td>
<td>Best intelligibility, fits well in the tempo of the song.</td>
<td>C</td>
<td>Comb filter effect, sounds amateur like, the vocals are separated from the rest of the mix.</td>
</tr>
<tr>
<td>Subject 2</td>
<td>E</td>
<td>The vocals and the rest of the song felt as a single piece. The vocals didn’t feel artificially added. The vocal fitted well in with the song.</td>
<td>C</td>
<td>The vocal felt a bit as a closet, the vocal did stand out from the rest of the stimuli’s, the vocals felt unnatural.</td>
</tr>
<tr>
<td>Subject 3</td>
<td>A</td>
<td>Natural, Adequate amount of space but still much intelligibility in the sound</td>
<td>D</td>
<td>Altogether to dry, felt as if the sound is sucked out of the ears.</td>
</tr>
<tr>
<td>Subject 4</td>
<td>A</td>
<td>Because it has an adequate amount of reverb at the same time as it has intelligibility and fits best in the mix.</td>
<td>C</td>
<td>Altogether to much reverb, (To large room) becomes somewhat indistinct. Dose not fit the mix and with the other instruments because of these reasons stated above.</td>
</tr>
<tr>
<td>Subject 5</td>
<td>A</td>
<td>I think that stimuli A has a good balance between the direct sound and the room sound/reverb, that’s why I have chosen this stimuli, it gives the vocals the intelligibility that I prefer.</td>
<td>D</td>
<td>The sound image becomes strange, the vocal ends up to close to the mix because the vocal is perceived very dry, it becomes an unbalance in the mix. It feels like the vocals are in one room and the other instrument are in a different room.</td>
</tr>
</tbody>
</table>
7.5 Conclusions from the listening test

How are each lead vocal perceived in each one of the mix?

- **Intelligibility**
  
  - The data indicates that the church have low and moderate intelligibility and could be due to the long decay time, which are the longest for all of the environments, this is can be seen in table 8 in the appendix, in the decay time measurements.
  
  - The data could be due to the controlled acoustical environment, which the song booth (K9) environment possesses, this are stated in the descriptions of the environment in the introduction of this paper.
  
  - The data indicates that the entrance hall is an environment with low intelligibility and this is due to the uncontrolled acoustical environment, perceived in the measurement and description of the entrance hall.
  
  - The data indicates that the anechoic chamber which possesses no reflections are the most intelligible environment. The data also indicates that there are some scattered opinions about the intelligibility of this environment.
  
  - The data indicates that song booth (K9) environment are perceived low and moderate in intelligibility, this could be due to the low level of the reverb which was added or the low diffusion and short pre delay time of the reverb.

It is indicated that there are different opinions amongst the subjects in which what environment is intelligible, however there are patterns that emerges from this ranking, which shows that the environments which have high intelligibility have low amount of reflections (the song booth (K9), anechoic chamber) and have controlled acoustics. It is also indicated that the environments which have reflections, the church, song booth (K9) and the entrance hall have low intelligibility, this is discussed in the analysis.

- **Perceived volume**
  
  - The data indicates that the church environment is perceived as high in volume due to the subjects ranking
  
  - These results indicate that the song booth is perceived as low in volume due to the ranking which shows that the environment is placed by three subjects in the low end of the scale.
  
  - The majority of subjects rated the entrance hall environment as low in volume as seen in the graph, however there are one subject that dose not agree with the other subjects, this could be due to different preference.
  
  - The data indicates that anechoic environment are perceived as loud, low and moderate in volume. The low and moderate are in majority, this could be due to the no reflections in this environment, but this is uncertain and need more investigations.
  
  - The data indicates that there are a majority of the subjects in which perceive the song booth (K9) + reverb environment as loud in volume, due to the subjects rating.
The data indicates on patterns that illustrates that there could be a relationship between the amount of reflections and perceived volume of the vocals, a higher amount of reflections the lower the volume are perceived and with lower amount of reflections the higher the amount volume becomes. However the environment with the least perceived volume is the entrance hall.

- **Amount of reverberation**
  - The data indicates that the church is perceived as reverberant but low in reverberation by the subjects.
  - The patter shown in song booth (K9) is mostly due to the subjects own preference and could be perceived as an anomaly.
  - The data clearly shows that the subjects perceive the entrance hall environment as the most reverberant environment. There are also a stray on the ranking of the stimuli by one subject, this are also an anomaly which indicates that the subjects have different preference in which they perceive sound.
  - The data from anechoic environment illustrates that the subjects perceives if an environment is non reverberant.
  - The data indicates also the church environment that it is an reverberant environment, however the song booth (K9) + reverb environment are perceived as more reverberant, due to the subjects ratings, then the church, this is clearly shown in figure 8, 3.

The rankings by the subjects indicate that the subjects can perceive changes in reverberation. The ranking also indicate that the subjects have own preferences when they evaluate reverberation due to the amount of strays, clearly show in the data, but there are similarities illustrated in the graph with a broad overview. The entrance hall is perceived as to have the most reverberation, the song booth (K9) + reverb have the second most amount of reverberation, the church in a third place, the song booth (K9) the fourth and the anechoic environment the least amount of reverberation.

**Perceived room size**

- The data indicates that the church is perceived as large in room size, but not the largest room.
- This illustrates that the song booth (K9) is perceived larger then the church in size but not the largest perceived room.
- The data clearly shows that the entrance hall is perceived as the largest room size.
- The data from anechoic chamber indicates that reflections could be a part of determine if a room is large or small in size.
- The data are indicating towards that the song booth (K9) + reverb environment is perceived as large in size, however three subjects stray from this, both towards a larger and a smaller room size, this could be due to the different perceptions of reflections by the subjects.
This data drawn from this question illustrates that there are a pattern in which the subjects uses the amount of reflections/reverberation to determine the size of rooms. With a large room size a lot of reflections/reverberation and a small room size a low amount of reflections/reverberation. An investigation in depth is done to see if there is a positive or negative affects of the perceived room size in each stimulus, which are presented in the following question.

Did the perceived room size affect the vocals positive or negative in the mix?
The pattern that is emerging in the data illustrates that depending on the type of reflections/perceived room size adds a positive or negative feature to the sound, there are a majority of the subjects that perceived the entrance hall as the largest room with the most amount of reverberation, however the entrance hall environment have the most negative affect on the sound, and the church which was perceived not as the largest room but attain the most positive features. There are also a pattern that indicates that non exciting reflections are a negative feature, however a low amount of reflections are perceived as both negative and positive, this is mostly due to the subjective liking from the subjects.

These results are very subjective answers; both the positive and negative motivations of the each environment are essentially preference dependent and are hard to draw conclusions from, however it is clearly stated that the entrance hall have no positive effect on the vocals and that the church environment have the most positive effect on the vocals versus the other environments.

Rank the 5 lead vocal environments from mostly preferred to least preferred by you in the mix.
From this data there are clearly indications that the entrance hall are the least preferred environment, this indicates as shown in the previous questions that it is ranked low by a majority of the subjects, why the subjects ranked this low are investigated in the next question. The graph indicates that the anechoic chamber is ranked low by the subjects, this could be due to the non-existent reflections which are also indicated in the previous questions. The song booth (K9) is perceived as a moderate environment by the subjects. All rankings of the song booth are between 5 -10 on the scale, this indicates that there is no major disliking on this environment. The song booth (K9) + reverb environment have extremes in the data, the majority of subjects ranked the environment over the middle of the scale (number 5 on the scale), this indicates that it is perceived as highly preferred, this is indicated in the previous question also, due to the amount of positive affects stated by the subjects of this environment.

The church environment is perceived as the most preferred environment. This is clearly shown in the graph, the are some stray data also, this could be due to the subjects liking, and why the subjects have rated this environment in this order will be investigated in following question.

Why is it most preferred, Least preferred
The church is the most preferred environment of the five environments because the adequate amount of reverb in this environment, and the intelligibility that the subjects perceive from this environment. Two of the subjects preferred the song booth (K9) and the song booth (K9) + reverb environment could be because they have different preference then the other subjects, so this indicates that there is no consistency in the data. However this suggests that there could be an advantage in using natural ambiance then post processed artificial ambiance on vocals. This is only an indication, however to be really certain, investigations on a more vast pool of subjects should be preformed, for statistically verify or falsify this indication. The data indicates that the least preferred environment was the entrance hall, this could be due to the uncontrolled acoustics and short pre delay times in this environment, this last place are shared with the anechoic environment which two subjects rated as their least preferred environment, this is due to the lack
of reflections which causes an unnatural feeling amongst the subjects, this indicates that there are could be an relationship between perceived “good” vocals and room ambience, this is also an side track for further investigation.

7.6 Analysis of the listening test

The findings from the listening test and analysed are discussed below.

How are each lead vocal perceived in each one of the mix?

• Intelligibility of the lyrics

All the stimuli’s except the entrance hall are ranked by the subjects as the most intelligible; this might be due to the fact that the entrance hall environment has a very reverberant acoustic which affects the sound in an unpleasant way perceived by the subjects. One environment has a larger room volume and the other environments have controlled acoustics which could eliminate early room reflection which could colorize the direct sound in a negative way [7]. This could be why the environments except for the entrance hall are preferred as highly intelligible.

Two subjects are choosing the anechoic environment; this could be due to fact that there are no early reflections from the room which could interfere with the direct sound [2].

Three subjects ranked the song booth (K9) environment as the environment with the second best intelligibility; this could be due to the controlled acoustic environment in the song booth. The room has damping on the walls and ceiling which gives control over the reflections, this was indicated in the description of the song booth environment and the decay time measurements conducted for this paper, however two subjects ranked the anechoic environment and the church the second most intelligible environment. This could be an effect of the missing reflections from the anechoic environment, and the long pre delay time in which the church have, see table 8 in the appendix and D. Griesinger paper [2].

All environments except the entrance hall was rated by the subjects a three on the scale which indicates that the subjects perceived the intelligibility of these environments as to be not high in intelligible or low in intelligibility.

The church, entrance hall and song booth (K9) + reverb environment was ranked as the second least intelligible environments, all these environments have audible reflections, this patten might indicate that the subjects who ranked the environments uses reverberation in someway to evaluate intelligibility.

The environment which was ranked the least intelligible environment by a majority of subjects was the entrance hall this could be due to lack of controlled acoustics which was perceived under the decay time and pre delay time measurements. This could be due to the early reflections interfering with the direct sound, which causes the low amount of intelligibility or clearness as D, Griesinger discusses in Room Impression, Reverberance, and warmth in Rooms and Halls [2].

• Perceived volume

The church and song booth (K9) + reverb were graded as the environments with the most perceived volume. The church and song booth (K9) + reverb was graded by two subjects
each, this could be due to the room ambience in which the volume of the vocals could be perceived as louder.

The church, song booth (K9), entrance hall and song booth (K9) + reverb were ranked as the seconds with most perceived volume, however there are two subjects how ranked the church as the one with the second most perceived volume, this might be due the long pre delay time or the long decay time in the environment, see tables 8-10 in the appendix, however the data are scattered.

The church, song booth (K9) and anechoic chamber were ranked a three on the scale, the anechoic chamber was ranked with the grade three by three subjects. The church and song booth (K9) were ranked three by one subject each, this indicates that the anechoic environment are perceived by a majority of subjects as the one with not the loudest or lowest volume.

The church, song booth (K9), entrance hall and song booth (K9) + reverb are graded as the environments with the second least perceived volume. The entrance hall and song booth (K9) were rated by two subjects each which indicates that environments with low in room reflections could be perceived as low in volume. However, the Song booth (K9) + reverb room perceived as low in volume, so the data are non consistent.

The song booth, entrance hall, anechoic chamber and song booth (K9) + reverb environments are perceived as the environments with the least perceived volume. Two subjects rated the entrance hall as the environment with the least perceived volume which indicates that the entrance hall have the lowest perceived volume of the environments.

- **Amount of reverberation**
  The entrance hall are perceived by four subjects as the environment with the most perceived amount of reverberation and one subject strayed from the rest of the group by choosing the song booth (K9) + reverb environment as the one with the most perceived amount of volume. This indicates that the entrance hall are perceived as the most reverberant room, this is most likely due to the un-acoustical processed reflective walls in the room and the short pre delay times measured in the room. The strayed data could be a result from that the subject has interpreted reverberation differently and/or have another preference to reverberation then the rest of the subjects have.

  The church, song booth and song booth (K9) + reverb are perceived as the second most reverberant environment. The church and song booth (K9) + reverb environments were rated by two subjects each and the song booth was rated by one subject. The data indicates that the environments which have long reverberation times are perceived as reverberant however there are some strays in the data.

  The church, song booth (K9) and song booth (K9) + reverb are perceived by the subjects as not low or high amount of reverberation, two subject’s rates the church and two subjects rate the song booth (K9) + reverb a three one the scale, this indicates that the subjects perceive these environments with an medium amount of reverberation. One subject rates the song booth (K9) as a three on the scale, which could be due to the controlled acoustical environment in the song booth which the decay time measurements states.
The church, song booth and entrance hall are perceived as the second least amount of reverberation by the subjects, three subjects have rated the song booth (K9) as the second least perceived amount of reverberation, this could also be due to the controlled acoustical environment in which the song booth are. However, two subjects rate the entrance hall and church as the second least perceived environment, these data could be due to different perceptions of reverberation by the subjects.

All the subjects rated the anechoic environment as the environment with the least amount of reverberation, this data are consistent and illustrates that when no reflections or no ambiance is present in the sound, its effects are noticeable.

**Perceived room size**

The church, entrance hall and song booth (K9) + reverb were ranked by the subjects as the largest rooms. Three of the subjects ranked the entrance hall as the largest room. This could be due to the amount of reflections in this environment and lack of controlled acoustics. The church and song booth (K9) + reverb environments have also reflections however these environments have longer decay times then the entrance hall, in which influence the subjects to perceive them as large rooms [7].

The song booth (K9), song booth (K9) + reverb were rated by the subjects as the second largest rooms. Three subjects rated the song booth (K9) as the second largest room and two rated the song booth (K9) + reverb the same. These data indicates that the song booth with controlled acoustics are perceived by more subjects then the controlled acoustical environment with artificial ambience, however the song booth (K9) + reverb data are scattered and one subject rated it as the largest room.

The church, song booth (K9), anechoic chamber and song booth (K9) + reverb are perceived as not large or small. Three subjects rated the church a three on the scale, one subject rated the song booth (K9) a three and one subject rated the song booth (K9) + reverb a three on the scale. This indicates that the church is the environment that gives the impression of being the medium sized room due to the subject’s ratings. No other pattern could be drawn.

The church, song booth (K9), anechoic chamber and song booth (K9) + reverb are the second smallest rooms perceived by the subjects, the pattern emerged from this data is that two subjects rated the anechoic chamber as the second smallest room, this could be due to the non existing reflections of this room. The subjects that rated other environments have chosen differently from each other and this makes the data non consistence.

The entrance hall and anechoic chamber are ranked as the smallest rooms, three rated the anechoic chamber as the smallest in size and two subjects rated the entrance hall as the smallest room. The subjects that rated the entrance hall could have been influenced by the short pre delay times and reflections of the room when they evaluated the perceived size. And the subjects how rated the anechoic chamber could have been influenced by the lack of reflections and there by rated it as the smallest size.
Did the perceived room size affect the vocals positive or negative in the mix?

The entrance hall have been perceived as the environment with the most negative affect by the room size, this is due to much reflections, to dry, sounds like a tin, to much room, gives an unclear and confused sound image and feels shut in, these parameter resembles the ranking of intelligibility, amount of reverberation and perceived volume. This could be due to the amount of uncontrolled reflections in the room, and are further examined in the next following questions.

The environment with the second most negative room size is the anechoic chamber. This was due to the too dry nature of the environment, no reflections from the environment and it is not perceived to fit the mix and music. This could be due to the lack of reflections from this environment. It indicates what happens to the perception of the vocals when the room ambience is gone from the recording.

The song booth (K9) was perceived as the environment with 2 negative and 2 positive impressions. The negative affect was natural, but to hard and to dry, this is probably due to the controlled acoustical environment in this room. The positive affects was that the room hade a good response, pleasant but not audible and the room size gave a pleasant space for the vocals. This clearly shows that subjects have different preferences when they evaluate sounds, due to the subjects explanations on the positive/negative affect of the room size.

The song booth (K9) + reverb had also both negative and positive affects, this environment hade a greater part of positive affects, it is farther back in the mix, this is caused by the spatial imaging, discussed in the background section. The environment have a good balance, a pleasant amount of room and reverb which indicates that low amount of reverberation subtly added to the sound are perceived as a positive trade from the listeners point of view, which was shown in the interview analysis. The environment are also perceived as not small and narrow, which indicates that the subjects is evaluating room size from the amount of reverberation. The negative aspects are that the environment was to big and to clear, this is could be due to the long pre delay time which was measured in this environment; with a longer pre delay time the reflection will not interfere with the direct sound [8] and can be a probable cause of this results.

The church is perceived as the environment with the most positive affects on the sound. This environment is perceived as it fitted the song, have an adequate tuning of the amount of reverb and adequate intelligibility and it sounds natural. This could be due to the decay time, which are longer then the decay time matured in the other environments, it has also the longest pre delay time of all environments. The negative affect perceived from this environment are; the reflections feels separated from the direct sound, the pre delay time is to long. This illustrates that this feature are perceived as negative and positive which are due to the subjects liking, however the majority of the subjects perceive this environment to be the most positive.
Rank the 5 lead vocal environments from mostly preferred to the least preferred by you in the mix.

The church, song booth (K9) and song booth (K9) + reverb are rated a 10 as the most preferred environments. The church are ranked by three subjects as the most preferred environment and the song booth (K9) and song booth (K9) + reverb environment are rated by one subject each. This could be due to the positive effects stated in the positive/negative question.

The song booth (K9) + reverb environment are rated a 9 of the most preferred which indicates that it is the second most preferred however there are only one subject how raked this way, which gives the data no constancy. The song booth (K9) + reverb environment are also rated by one subject a 8.5 on the scale which indicates what the environment are the third most preferred environment but this does not show any consistency.

The Church, song booth (K9) are rated a 7.5 on the scale by one subject each, this indicates that these environments are the forth most preferred environments.

The song booth (K9) and the anechoic chamber and song booth (K9) + reverb are ranked a 6.5 on the scale by one subject each and are placed fifth most preferred on the scale, this are also no consistency due to the subjects ranking.

The song booth (K9) are rated a 5.5 by one subject, this indicates that the environment is the sixth most preferred environment, this data are non consistent.

The song booth are rated 4.5 on the preference scale by one subject which makes it the seventh most preferred environment.

The church and the anechoic chamber are rated a 3.5 on the scale which indicates that they are the eights most preferred environment; each one are rated by one subject each, which gives the result no consistency.

The entrance hall are rated 2,5 on the scale by one subject, and are the ninth most preferred environment, which gives no consistency in the data.

The song booth (K9) + reverb and the entrance hall are the tenth most preferred environments rated by one subjects each, which are not consistence in the data.

The entrance hall and anechoic chamber are the least preferred environments, three subjects rated the entrance hall as the least preferred and the two subjects rated the anechoic chamber as the least preferred environment. This result could be due to the opinions which were stated in the previous question, however why the subjects ranked the stimuli’s in this order will be examined in the next question.

Why is it most preferred, Least preferred

The church are perceived as the mostly preferred because it attains an adequate amount of reverberation and are perceived as intelligible and corresponds to the mix, these features are also show in the ranking questions. For the intelligibility question two subjects ranked it high and one who preferred this environment, ranked it with low intelligibility, this indicates that the mostly preferred environment do not have to have a high amount of intelligibility. The subjects preferred this environment also because of the balance between the direct sound and the reflections, this is
indicated in the ranking of the amount of reverberation, were the church environment are perceived reverberant but with a low amount.

One subject rated the song booth (K9) as the preferred environment, because it fitted well with the tempo of the song, and had good intelligibility, which also shows in the subjects ranking of intelligibility were the subjects ranked the song booth (K9) as the most intelligible environment.

One subject also rated the song booth (K9) + reverb as the most preferred environment, the vocals and the rest of the song felt as a single piece, the vocals didn’t feel artificially added and the vocal fitted well in with the song, the subjects rated the environment as low in intelligibility, high in volume and large in reverberation, this indicates that the subject does not perceive high intelligibility as an factor in which relates with their preference.

Three subjects did prefer the entrance hall least, due to perceived comb filter effects, separation between the vocals sound and instruments sound, the vocals have an unnatural feeling, too large room becomes indistinct and are perceived as amateur like and sound as a closet. The comb filter effects are could be due to the short early reflections show in the pre delay time measurements, see table 11 in the appendix. In this case the short early reflections have colored the sound negatively [7] and could be the major cause of why the subjects perceived the vocal sound to be out of place in the mix, indistinct and unnatural in nature. This could also be a parameter which interfered with the perception of intelligibility and an answer to why the subjects ranked the amount of intelligibility as low in this environment and why the subjects felt negative affects on the room size.

Two other subjects rated the anechoic chamber as the least preferred environment, the environment was perceived as to dry, the sound becomes unbalanced in the mix, are perceived as been recorded in another room from the instruments, and are perceived as to close to the listener in the spatial image. The dry fullness of environment are in direct link with the extreme controlled acoustics in the anechoic chamber. This gives the subjects the perception of the environment is to dry. The unbalance and external recording effect perceived are also liked with the acoustical properties of the environment. The instruments are recorded in a controlled acoustical environment but with a sound of the room in which the anechoic chamber lacks, this is mostly the cause of the unbalanced and the different recording room feeling perceived by the subjects. The perception of the closeness of the vocal to the listener are due to how humans perceive sound sources in the sound image, which are discussed in the background section earlier in this paper and could be a reason of why the vocal are perceived as to close to the listener, this is indicated also in the question of perceived volume when the subjects rated the vocal stimuli a three on the scale, which indicates that the environment are in the middle of to high and to low. The ranking questions are also indicating that the sound are perceived by a majority of subjects as dry, high in intelligibility, low in reverberation and small in room size, which is also indicated here in this question.

7.7 Discussion of the listening test
There were no clear indications which showed that there are an advantage in using real room ambience, scattered answers, over post processed artificial ambience. It only shows small patterns, from this listening test. The patterns show that there could be an advantage, and that the church environment could be the most effective and the entrance hall the least effective. Although the second most preferred environment was the post processed which makes the pattern vague. All patterns and indications need further investigation. The subjects where consistence in there answers and ranking thru the listening test, as explained later in the discussion, however these
patterns needs to be examined further in detail, pattern by pattern because of the vast subject this thesis discusses.

The intelligibility data indicates that there is no consistency in what environment are perceived as the most intelligible one, due to the low amount of listening test subjects. However there are a pattern in which it indicates that there are an relationship between perceived intelligibility and room reflections, the more audible the reflections are the less intelligible the vocals become, further investigations in to this is required, it will not be looked into in this paper. There are patterns in the perception of volume data which indicates that environments which are rich in reflections versus the environments with low amount of reflections are perceived as louder in volume, in spite of some strays in the subjects ranking. This pattern can be closer investigated by expanding this listening test to a more vast audience, to see if this is a pattern that holds in an expansion of listening test subjects or is it just occurring it this particular event. There are clearly patterns in the amount of reverberation data in which the amount reverberation could be perceived. It also indicates that the subjects can perceive reverberation from no existing reflections and vice versa. The data illustrates that the perception of amount of reverberation between environments are vague but a pattern can be shown, however further work would be needed for getting a clear result of these data.

The affects on the environments attains booth negative and positive features for a majority of environments, shown in the analysis of the listening test section, however the data from the positive negative question supports the ranking questions:

Example; The entrance hall are ranked as the environment as high in reverberation, and low in intelligibility, and the negative features stated are to much echo, to much room, confused sound image and are least preferred by a majority of subjects.

These patterns are indicated in all environments, however they are not examined further in this paper, and can be interesting further work.

The four major patterns that are showed in perceived room size data is that the entrance hall is perceived as the room with the largest room size, the song booth (K9) as the second largest room, the church as the room with the medium room size due to the subjects ranking, and the anechoic chamber with the smallest size. However, the data are clearly scattered, there are two subjects in which rated the song booth (K9) + reverb as the second largest room and one subject in which rated the entrance hall as the smallest sized room. These patterns need further investigation, to really understand how room sizes are perceived. The individual results have no consistency, but the pattern that emerges shows that the church, song booth (K9) + reverb and the song booth have their ratings in the upper end of the scale which indicates that they are mostly preferred by the subjects versus the anechoic chamber and the entrance hall which have their ratings in the low part of the scale, see figure 9 in the appendix for a broad overview of the pattern.

7.8 Interesting findings of the listening test

This section summarizes the interesting findings of the data, conclusions, analysis and discussion.

How are each lead vocal perceived in each one of the mix?

The entrance hall is perceived by a majority of subjects, as the environment with the least intelligibility of all the stimuli’s and the church are perceived as the most intelligible environment. There are three environments that are perceived high in volume by the subjects; the church, song booth (K9) + reverb and anechoic environment. The entrance hall are rated as low in volume by two subjects, however there are scattered data which shows that the song booth (K9), anechoic
chamber and song booth (K9) are low in volume. The amount of reverberation perceived by the subjects shows that the entrance hall are perceived with the most amount of reverberation, then the song booth (K9) + reverb and the church as the environments with the second most reverberation. The song booth (K9) is the environment with the second least reverberation and last the anechoic chamber with the least amount of reverberation.

**Perceived room size**
The trend clearly show that the entrance hall is perceived as the largest room, the song booth as the second largest room, the song booth (K9) + reverb on third place and forth place the church and last the anechoic chamber as the smallest in size. The church is not perceived as low reverberant although it is on third place. It is also a pattern that indicates that there are a relationship between reflections and perceived room size.

**Did the perceived room size affect the vocals positive or negative in the mix?**
Four subjects perceived positive effects of the room size from the church environment because the vocal fitted in the song, sounded natural, had adequate tuning of the amount of reverb and adequate intelligibility. Three subjects perceived positive affect from the room size on the song booth (K9) + reverb environment, it is dose not sound small/narrow, have a pleasant amount of room for the type of music, have a good balance amount of reverb, the vocal ends up farther back in the mix, and has a good balance between the vocals and instruments. Four subjects perceived negative effects of the room size of the anechoic environment, does not feel as a part in the song, clear but to dry, ends up to close to the listener in the mix, hard to understand the vocals. All subjects perceived negative effects of the room size from the entrance hall environment because of the sound sounds shut in, assertive, have indications off comb filter effects, to dry, attain to much echo, sound like a tin, to much room can be unclear and give arise to a confused sound image.

**Rank the 5 lead vocal environments from mostly preferred to the least preferred by you in the mix.**
The most preferred environment is the church, due to adequate amount of reverb, but still has much intelligibility in the sound and fits best in the mix. The least preferred environment are the entrance hall, due to comb filter effects, sounds armature like, are perceived as separated from the rest of the mix, feel like in a closet and felt unnatural, to much reverb becomes some what indistinct, do not fit the mix. The subjects ranked the anechoic environment as the almost least preferred environment due to dry nature of the environment, the sound is sucked out of the ears, the sound image become strange, are perceived as to close in the mix and becomes unbalanced and is perceived as it is in a different room then the instruments.

**8. Discussion on the work, methodology and future work**
In this section are the general discussion about the work processes and the methodology of the paper is presented.

**Discussion concerning the background**
A literature study on how humans perceive the general reflections and reflections in room environments could have provided deeper understanding of why the subjects evaluated and ranked the environments as they did in the listening test. These results and further results will need to be compared to findings in this relevant area.
The measurements conducted in the vocal environments are only on decay time of each environment with a speaker with limited frequency range, with a maximum of three
measurements position in the largest environment and only one in the smallest room. These measurements could have been more detailed with more measurement positions, with a full range speaker as sound source in each environment and attaining the absorption coefficient on the walls, ceilings and floors to specify the amount of reflections in each environment.

**Discussion concerning the interview analysis**

Instead of conducting interviews with few mixing engineers to determine how to mix the stimuli for the listening test, the book, *the mixing engineer’s handbook* written by Bobby Owsinski, was chosen. This book compiled prior interviews with 20 professional mixing engineers in which 11 attained vast experience in mixing low-dynamic music, and where chosen on these premises. This gives arise to a more vast selection of opinions on how to mix low-dynamic music, but limits the data points that could be drawn from interviews due to the already made interviews. The questions for the analysis were broken down to five questions with seven sub questions for getting manageable data from the interviews. This gave rise to a large amount of information, but limited data points due to the fact that none of the engineers were asked the questions that was stated for this paper. This gave scattered data responses from two – eleven data points, from each question/sub question. This data gave an insight in to how the engineers worked with low-dynamic music and what effects were used. It could have been beneficial for the project if the researcher for this project have conducted own interviews. This would have given more coherent interview data from the stated questions, and attained more qualitative data. The results stated from the interview analysis are patterns, clues and tips on how to conduct mixing of low-dynamic music.

The mixing constraints for the stimuli are based on the results from the interview analysis and the literature review in the background section, on how each effect were going to be used, in the mixing process. Thus, constraints could have been more methodical if the own interviews designed for this experiment had been conducted due to the more precise data that could have been attained from the interviewed engineers instead of patterns, clues and tips on how to mix low-dynamic music. These mixing constraints could have also stated by only doing a deeper literature investigation, however the interviews gives a more direct approach on how engineers mix low-dynamic music, then only a deep literature review could have attained. But, these published interviews did provide interesting references for common mixing practice and a way to frame this exploratory study. The volume levels of the instruments and vocals in the mix were set by the researcher/author by his own perception of low-dynamic. This could have been done by a professional mixing engineer with great experience of low-dynamic music mixing. This was not an option due to narrow time frames of the project and the knowledge attained by the researcher of the interview analysis and what constraints were going to be used in the mix.

**Discussions concerning the listening test**

The listening tests purpose was to give an answer to the research question/hypothesis stated in the introduction of the paper, by letting five subjects with audio engineering education evaluate the five vocal environment in low-dynamic music and give their opinions in words on each environment. Only five subjects were chosen for this work due to time limitations of the work and for getting manageable data, and see if there could be an advantage of using room ambience.

The questions were to examine how the subjects evaluated different parameters of the environments. The ranking questions did show how the subjects perceived each environment to the parameters set. The questions with motivational answers gave a clear view over how the subjects perceived each individual environment. The cons of the questions 1 and 2 were that they didn’t show was the most or least preferred environment and why.

One solution is to have more questions about why the subjects perceived the environments in the order they have chosen. This could shine more light on the perception of vocal environments in
low-dynamic music. A statistical approach to the listening test analysis was not chosen because of the small number of subjects for the listening test and the qualitative approach of the questions, and to see there would be interesting to make a larger scale listening test based on this work. The data from the parameters and ranking questions were to some extent scattered, conversely in the motivation part of the questions, were the answers more consistent, but they indicated only patterns. The amount of subjects was also insufficient when rendering the data. The data only indicates only patterns and a number of consistencies of the subjects liking. It was not possible to group subjects into types of listeners with similar preferences or perceptions. A greater amount of subjects could have been in order for this test, to statistically verify or falsify these patterns which are illustrated in the data, and there by getting a greater view over the evaluated environments and to state if there is an advantage in natural ambience on vocals. The expected results from the listening test were indicated to be correct from the results drawn from the data; it is only indications, where further work would be necessary to state these results as facts. The interface was programmed to give the subjects full control over the playback of the stimuli. This gives the subjects free hands in listening to the environments. A problem arose in which the buttons on the interface stop working in the middle of the test for each subject. This could have been prevented by the researcher, with conducting more then one evaluation test of the interface prior to the listening test and thereby eliminate the problem. After the test when asking the subjects if restarting the interface was an obstacle in which influenced them negative, some subjects perceived it as minor discomfort. This is could be a source of error in the listening test.

9. Conclusions of the work

The initial question for this work was to see and understand if there were an advantage in using natural ambience on vocals instead of using post processed artificial ambience on the vocals in low-dynamic music. The researcher/author attempted to answer this question by evaluating five different vocal environments in low-dynamic music. To obtain a clear overview and acquire clear data and to answer this question the work was broken down in to three parts, background, interview analysis, and listening test.

Interview analysis
The interview analysis showed that the engineers used equalizers, to cut or boost frequencies, which was also stated in literature review, and the analysis showed that some engineers used frequency starting point when mixing, because of experience and knowledge of what type of instrument or voice type lies in special frequency bands. A majority of engineers used compression only in the mixing stage, and a third of the engineers used compression on everything, this gives an hint that compression are effective in low-dynamic music for controlling sounds, which also the literature review stated. A majority of engineers implement the compressor over a bus and then added it to the initial sound, and it was indicated that the compressor was used as an effect more then a dynamic tool. The delay effect is being used to a larger degree of the engineers, and mostly on vocals often with the tempo of the song, a majority of the engineers used also reverbs mostly on vocals. In the analysis it was indicated that the use was to give the vocals the right character and that the plate reverb were the most preferred reverb unit amongst the engineers, why this is, was unclear. However two of the engineers worked on the recorded ambient (or ambient microphones) which followed every recorded instrument, before they reconsidered the use of reverb effects, this was to add first reflections to the sound.
**Listening test**
The entrance hall are perceived as the least intelligible, perceived as low in volume, attain the largest room size, have the most negative features, have the most amount of volume and are least preferred by a majority of subjects, due to comb filter effects, sounding armature like, are perceived as separated from the rest of the mix, feels like in a closet, felt unnatural, to much reverb becomes some what indistinct and does not fit the mix. The perceived negative affects of the room size from the entrance hall environment; the sound sounds, shut in, assertive, have indications off comb filter effects, to dry, attain to much echo, sound like a tin, to much room can be unclear and give arise to a confused sound image. The church is perceived as the most intelligible environment, high in volume, it is not perceived low in reverberation, and is the most preferred environment due to adequate amount of reverb but still have much intelligibility in the sound and fits best in the mix.

Perceived positive effects of the room size from the church environment; the vocal fitted in the song, sounded natural, and had adequate tuning of the amount of reverb and adequate intelligibility.

There are an amount of patterns and indications in the data from this work in this thesis, in which all needs further work for understanding.

- The listening test indicates patterns that there are a relationship between reflections and perceived room size due to the subjects ranking of perceived room size.

- The intelligibility ranking indicates that there is a relationship between perceived intelligibility and room reflections, the more audible the reflections are the less intelligible the vocals becomes.

- There are also patterns in this data which indicates that environments which are rich in reflections versus the environments with low amount of reflections are perceived as louder in volume.

- There are patterns in these data in which the amount reverberation could be perceived and that the subjects can perceive reverberation from no existing reflections and vice versa. The data illustrates that the perception of amount of reverberation between environments, the data are vague but a pattern seems to emerge. It requires further testing.

- The question of which environment is most/least preferred illustrate that these environment can be shown, however the other results from the rating have no consistency, but the pattern that emerges shows that the church, song booth (K9) + reverb and the song booth have their ratings in the upper end of the scale which indicates that they are mostly preferred by the subjects versus the anechoic chamber and the entrance hall which have their ratings in the low part of the scale.

The studied shows that are indications that the natural ambiance on vocals could be an advantage in low-dynamic music due to the subjects ratings from the preferences scale, nonetheless further work would be needed to state this as a fact. The studies handles a large topic and work could have been divided in to two separate research questions, how to mixing engineers mix low-dynamic music, how do humans perceive vocals in low-dynamic music, for attaining deeper understanding in the those areas which are studied in the paper. A more statistical study could be adequate as a follow up study on these patterns, indications shown in this paper.
This topic gives many indications and shows a lot of patterns which would be interesting to investigate further to really get a clear result if it is possible or not possible to use real vocal ambience versus post processed ambience in low-dynamic music. Further investigations can give an understanding on how the listener perceives vocals in low-dynamic music, and could give the engineer which mixes low-dynamic music a guideline on how to mix the vocals, and the result can also give the engineer a more objective approach instead of many subjective opinions from listeners, and give the listener a better experience when listening to low-dynamic music.

The researcher has learned a great deal of how to use effects on vocals from the interview analysis, how different room ambience can alter the vocal sound negatively and positively, and to an extent how people perceives the environments. It has also given a solid base of information on patterns to follow and conduct further investigation on.
10. References

Boston, MA, USA: Course Technology, Incorporated, 1997
ISBN 1-918371-17-1

[2] Griesinger, David; *Room Impression, Reverberance, and warmth in Rooms and Halls,* Presented at the 93rd AES Convention, 1992 October 1-4, San Francisco

USA, TAB Books, 1994


Vallejo, CA, USA: Artistpro.com, 1999

California, USA: Waldsworth, Inc., 1994
ISBN 0-534-17646-1

[7] Rumsey, Francis; *Spatial Audio*
ISBN 0-240-51623-0

[8] Griesinger, David; *Spaciousness and envelopment in musical acoustics,*
Presented as the 101st AES Convention, 1996 November 8-11, California.
11. Appendix

Microphone and Microphone pre amplifier adjustments

Table 7

<table>
<thead>
<tr>
<th>Adjustments</th>
<th>Microphone pre amplifier</th>
<th>Microphone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>Summit Audio 2BA-221 Mic and Line module</td>
<td>AKG c414</td>
</tr>
<tr>
<td>Hp filter</td>
<td>200 Hz</td>
<td>150 Hz</td>
</tr>
<tr>
<td>Characteristics</td>
<td></td>
<td>Cardioid</td>
</tr>
<tr>
<td>Damping/pad</td>
<td>0</td>
<td>- 10 dB</td>
</tr>
<tr>
<td>Mic Gain</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Polarity</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>+ 48 V</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Tube output</td>
<td>1.5</td>
<td></td>
</tr>
</tbody>
</table>

Measurement of decay time

Table 8

<table>
<thead>
<tr>
<th>Church</th>
<th>Measured frequencies Hz</th>
<th>Decay time seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mic position</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31,5</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>63</td>
<td>0.447</td>
<td></td>
</tr>
<tr>
<td>125</td>
<td>0.459</td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>0.504</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>1.076</td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>0.974</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>1.111</td>
<td></td>
</tr>
<tr>
<td>4000</td>
<td>0.606</td>
<td></td>
</tr>
<tr>
<td>8000</td>
<td>0.365</td>
<td></td>
</tr>
<tr>
<td>Position b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right of the altar</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31,5</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>63</td>
<td>1.224</td>
<td></td>
</tr>
<tr>
<td>125</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>1.391</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>1.451</td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>2.312</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>1.721</td>
<td></td>
</tr>
<tr>
<td>4000</td>
<td>1.393</td>
<td></td>
</tr>
<tr>
<td>8000</td>
<td>0.852</td>
<td></td>
</tr>
<tr>
<td>Position c</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left back corner</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31,5</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>63</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>125</td>
<td>0.597</td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>1.727</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>1.523</td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>2.241</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>1.713</td>
<td></td>
</tr>
<tr>
<td>4000</td>
<td>1.231</td>
<td></td>
</tr>
<tr>
<td>8000</td>
<td>0.784</td>
<td></td>
</tr>
</tbody>
</table>
### Table 9

<table>
<thead>
<tr>
<th>Song Booth K9</th>
<th>Measured frequencies Hz</th>
<th>Decay time seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position a</td>
<td>Mic position</td>
<td></td>
</tr>
<tr>
<td>31.5</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>63</td>
<td>0.352</td>
<td></td>
</tr>
<tr>
<td>125</td>
<td>0.447</td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>0.432</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>0.378</td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>0.326</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>0.316</td>
<td></td>
</tr>
<tr>
<td>4000</td>
<td>0.349</td>
<td></td>
</tr>
<tr>
<td>8000</td>
<td>0.373</td>
<td></td>
</tr>
</tbody>
</table>

| Position b    | Middle of room          |                    |
| 31.5          | 0                       |
| 63            | 0.355                   |
| 125           | 0.44                    |
| 250           | 0.39                    |
| 500           | 0.398                   |
| 1000          | 0.376                   |
| 2000          | 0.375                   |
| 4000          | 0.374                   |
| 8000          | 0.385                   |

### Table 10

<table>
<thead>
<tr>
<th>Entrance Hall</th>
<th>Measured frequencies Hz</th>
<th>Decay time seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position b</td>
<td>Far end of hallway</td>
<td></td>
</tr>
<tr>
<td>31.5</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>63</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>125</td>
<td>0.7</td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>0.463</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>0.423</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>0.375</td>
<td></td>
</tr>
<tr>
<td>4000</td>
<td>0.387</td>
<td></td>
</tr>
<tr>
<td>8000</td>
<td>0.359</td>
<td></td>
</tr>
</tbody>
</table>

### Table 11

<table>
<thead>
<tr>
<th>Reflection/Pre delay time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre delay time seconds</td>
</tr>
<tr>
<td>Angle from sound source</td>
</tr>
<tr>
<td>Vertical 0</td>
</tr>
<tr>
<td>Vertical 45</td>
</tr>
<tr>
<td>Horizontal 50</td>
</tr>
<tr>
<td>Interview analysis table</td>
</tr>
<tr>
<td>--------------------------</td>
</tr>
<tr>
<td><strong>Does the engineer work with equalizers?</strong></td>
</tr>
<tr>
<td><strong>Joe Chiccarelli</strong></td>
</tr>
<tr>
<td><strong>Lee DeCarlo</strong></td>
</tr>
<tr>
<td><strong>Benny Faccone</strong></td>
</tr>
<tr>
<td><strong>Jerry Finn</strong></td>
</tr>
<tr>
<td><strong>Andy Johns</strong></td>
</tr>
<tr>
<td><strong>Kevin Killen</strong></td>
</tr>
<tr>
<td><strong>Allen Sides</strong></td>
</tr>
<tr>
<td><strong>Don Smith</strong></td>
</tr>
<tr>
<td>Name</td>
</tr>
<tr>
<td>-----------------------</td>
</tr>
<tr>
<td>Guy Snider</td>
</tr>
<tr>
<td>Ed Stasium</td>
</tr>
<tr>
<td>Bruce Swedish</td>
</tr>
<tr>
<td>Joe Chiccarelli</td>
</tr>
<tr>
<td>Lee DeCarlo</td>
</tr>
<tr>
<td>Benny Faccone</td>
</tr>
<tr>
<td>Jerry Finn</td>
</tr>
<tr>
<td>Andy Johns</td>
</tr>
<tr>
<td>Kevin Killen</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Joe Chiccarelli</td>
<td>Yes</td>
<td>Mixing</td>
<td>For getting drums out of the way so that the vocals fit.</td>
<td>Less</td>
<td>Drums</td>
<td>Compressor over a stereo bus.</td>
<td></td>
</tr>
<tr>
<td>Lee DeCarlo</td>
<td>No</td>
<td>Mixing</td>
<td></td>
<td></td>
<td></td>
<td>Using limiters and gates on the sound</td>
<td></td>
</tr>
<tr>
<td>Benny Faccone</td>
<td>Yes</td>
<td>Mixing</td>
<td>Too keeping it controlled, smooth and consistent</td>
<td>Much</td>
<td>So its breathing right</td>
<td>Compress everything individually</td>
<td></td>
</tr>
<tr>
<td>Jerry Finn</td>
<td>Yes</td>
<td>Mixing</td>
<td>The artist and industry demands that</td>
<td>Much</td>
<td>For adding character to the sound.</td>
<td>Compressor over a bus (mult).</td>
<td></td>
</tr>
<tr>
<td>Andy Johns</td>
<td>Yes</td>
<td>Mixing</td>
<td>Uses it for modifying a sound.</td>
<td>Much</td>
<td>Everythi ng</td>
<td>Compressor on a bus (mult).</td>
<td></td>
</tr>
<tr>
<td>Kevin Killen</td>
<td>Yes</td>
<td>Mixing</td>
<td>It will be added later on at a later point during mastering and broadcast</td>
<td>Less</td>
<td>For adding character to the instruments</td>
<td>Compressor on a bus</td>
<td></td>
</tr>
<tr>
<td>Allen Sides</td>
<td>Yes</td>
<td>Mixing</td>
<td>For adding sustain without killing the attack, adding more ambient to the initial sound</td>
<td>Much</td>
<td>Compressor on a bus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>------</td>
<td>--------</td>
<td>-----------------------------------------------------------------------------------------</td>
<td>------</td>
<td>---------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Don Smith</td>
<td>Yes</td>
<td>Mixing</td>
<td>Making sounds push through the track</td>
<td>Much</td>
<td>Don’t have any consistence way of using it.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guy Snider</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Over empathized compressor on a bus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ed Stasium</td>
<td>Yes</td>
<td>Mixing</td>
<td>It adds tones to the instrument.</td>
<td>Much</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bruce Swedish</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Joe Chiccarelli</td>
<td>Yes</td>
<td>Mixing</td>
<td>The space between the initial sound and the echo makes the sound big as possible</td>
<td>Much</td>
<td>The space between the initial sound and the echo makes the sound big as possible</td>
<td>Vocals</td>
<td>Delayed echo</td>
</tr>
<tr>
<td>Lee DeCarlo</td>
<td>Yes</td>
<td>Mixing</td>
<td>Making the vocal stand out from the arrangement.</td>
<td>Much</td>
<td>Making the vocal stand out from the arrangement.</td>
<td>Vocals</td>
<td>Delay in the tempo of the song</td>
</tr>
<tr>
<td>Benny Faccone</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jerry Finn</td>
<td>Yes</td>
<td>Mixing</td>
<td></td>
<td>Much</td>
<td></td>
<td>Vocals</td>
<td>Tape machine or space echo Delaying the send to the plate with the tempo of the song.</td>
</tr>
<tr>
<td>Andy Johns</td>
<td>Yes</td>
<td>Mixing</td>
<td></td>
<td>Much</td>
<td></td>
<td>Vocals</td>
<td>A short 25-32 ms delay in the</td>
</tr>
<tr>
<td>---------------------</td>
<td>---------</td>
<td>------------------------</td>
<td>--------------------------------------------------------------</td>
<td>------------------</td>
<td>-------------------------------</td>
<td>----------</td>
<td>------------------</td>
</tr>
<tr>
<td>Kevin Killen</td>
<td>Yes</td>
<td>Mixing</td>
<td>Much</td>
<td></td>
<td></td>
<td></td>
<td>back of the vocal, for mid tempo songs longer delay.</td>
</tr>
<tr>
<td>Allen Sides</td>
<td>Yes</td>
<td>Mixing</td>
<td>Pre delay to the chamber</td>
<td>Less</td>
<td></td>
<td></td>
<td>Tape slap machine with varispeed</td>
</tr>
<tr>
<td>Don Smith</td>
<td>Yes</td>
<td>Mixing</td>
<td>The vocal becomes duller but fatter.</td>
<td>Much</td>
<td>The vocal becomes duller but fatter.</td>
<td>Vocals</td>
<td>Ampex mastering preview DDL</td>
</tr>
<tr>
<td>Guy Snider</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ed Stasium</td>
<td>Yes</td>
<td>Mixing</td>
<td></td>
<td>Much</td>
<td>Vocals</td>
<td></td>
<td>Slap tape in the tempo of the song, often with the return to a reverb chamber.</td>
</tr>
<tr>
<td>Bruce Swedish</td>
<td>Yes</td>
<td>Mixing</td>
<td></td>
<td>Much</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joe Chiccarelli</td>
<td>Yes</td>
<td>Mixing</td>
<td>Concentrate on giving the vocal the right character by trying to find effects which complement the voice better.</td>
<td>Much depending on the project</td>
<td>Adding personality to the sound</td>
<td></td>
<td>EMT plate, live chamber</td>
</tr>
<tr>
<td>Lee DeCarlo</td>
<td>Yes</td>
<td>Mixing</td>
<td></td>
<td>Much</td>
<td></td>
<td></td>
<td>Using different echo chambers</td>
</tr>
<tr>
<td>Benny Faccone</td>
<td>Yes</td>
<td>Mixing</td>
<td>Much and less depending on the artist and producer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jerry Finn</td>
<td>Yes</td>
<td>Mixing</td>
<td>Less</td>
<td>Tying too get</td>
<td></td>
<td></td>
<td>Plate reverb</td>
</tr>
</tbody>
</table>

69
<table>
<thead>
<tr>
<th>Name</th>
<th>Work</th>
<th>Sound</th>
<th>Amount</th>
<th>Room Ambiance</th>
<th>Reverb</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andy Johns</td>
<td>Yes</td>
<td>Mixing</td>
<td>Much</td>
<td>The dry sound</td>
<td>Vocals</td>
<td>Plate EMT 140</td>
</tr>
<tr>
<td>Kevin Killen</td>
<td>Yes</td>
<td>Mixing</td>
<td>Depends on the situation</td>
<td>Wants it to sound lush</td>
<td>Vocals</td>
<td>RMX 16 (non-line reverb)</td>
</tr>
<tr>
<td>Allen Sides</td>
<td>Yes</td>
<td>Mixing</td>
<td>Much</td>
<td>Trying to get a room sound, so it sounds natural</td>
<td>Vocals</td>
<td>RMX 16 (non-line reverb)</td>
</tr>
<tr>
<td>Joe Chiccarelli</td>
<td>Yes</td>
<td>Mixing</td>
<td>Going through a phase with less reverb</td>
<td>Less</td>
<td>Chamber, EMT140s</td>
<td>EMT250, EMT 252</td>
</tr>
<tr>
<td>Ed Stasium</td>
<td>Yes</td>
<td>Mixing</td>
<td>Goes through a phase with less reverb</td>
<td>Less</td>
<td>Chamber, EMT140s</td>
<td>EMT250, EMT 252</td>
</tr>
<tr>
<td>Swedish</td>
<td>Yes</td>
<td>Mixing</td>
<td>Going through a phase with less reverb</td>
<td>Less</td>
<td>Chamber, EMT140s</td>
<td>EMT250, EMT 252</td>
</tr>
</tbody>
</table>

The table is organized in rows, with the names of various recording engineers and their responses to questions regarding their work with natural ambience, under recording/mixing, and the desired sound qualities. Each row includes the name of the engineer, their work with natural ambience, the sound they desire, and the amount they prefer. The table also includes columns for room ambience, reverb, and specific types of reverb used. The final column provides details on what kind of sound or ambience is used in the recording process.
<table>
<thead>
<tr>
<th>Bruce Swedish</th>
<th>Yes</th>
<th>Mixing and recording</th>
<th>Too add first reflections to sounds.</th>
<th>Much</th>
<th>Instruments</th>
<th>First reflections</th>
</tr>
</thead>
</table>

**Figure 9: Broad view over the preference data**
Table 13

<table>
<thead>
<tr>
<th>Recording equipment</th>
<th>Mixing equipment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mac computer</td>
<td>Mac computer</td>
</tr>
<tr>
<td>Digidesign Control 24</td>
<td>PowerMac G5</td>
</tr>
<tr>
<td>Acer Aspire 3500</td>
<td>Double 1,8GHz PowerPC processor</td>
</tr>
<tr>
<td>E-MU 1616 Carbbus + Breakout box</td>
<td>1 GB DDR SDRAM memory</td>
</tr>
<tr>
<td>Pro Tools</td>
<td>Pro Tools 6.4</td>
</tr>
<tr>
<td>Cubase SE 3</td>
<td>Pro Tools standard plug ins</td>
</tr>
<tr>
<td>Audio Technica Headphones</td>
<td>Compressor</td>
</tr>
<tr>
<td>Genilec Monitors</td>
<td>1 band Eq</td>
</tr>
<tr>
<td>Pevey Guitar amplifier</td>
<td>4 band Eq</td>
</tr>
<tr>
<td>Yamaha Drum set</td>
<td>D-verb</td>
</tr>
<tr>
<td>Trace Elliot bass amplifier</td>
<td></td>
</tr>
<tr>
<td>Piano</td>
<td></td>
</tr>
<tr>
<td>Summit audio inc, 2BA-221 mic, line module</td>
<td></td>
</tr>
</tbody>
</table>
Listening test

Instructions for the test
This listening test is an evaluation of different lead vocal environments in one low-dynamic music arrangement and only the vocal sound/environment will be evaluated. Read the form before the test for getting a comprehension on what will be asked, and respond to each question with sincerity and well thought out answers.

Do you have any form of audio technical education?
Yes
No

(If yes) what education? __________________________________________________

Do you have any hearing disorder that you know about?
Yes
No

(If yes) what kind of hearing disorder? _______________________________________
____________________________________

How are each lead vocal perceived in each one of the mix?
Rank the stimuli’s and write the letter of the stimuli beside the corresponding number on the scale, a one is perceived as the least amount and five is the most amount perceived.

Intelligibility (clearness) of the lyrics

5 ______  Most intelligibly
4 ______
3 ______
2 ______
1 ______  Least intelligibly
Volume of the vocals, which is perceived loud or low versus the music arrangement.

5_______ Most perceived volume
4_______
3_______
2_______
1_______ Least perceived volume

The amount of reverberation sound of the vocals versus the direct sound of the vocals
i.e. which stimuli has what amount of reverberation ratio

5_______ Most Reverberation
4_______
3_______
2_______
1_______ Least Reverberation

Rank each one of the stimuli’s in order of perceived room size

5_______ Largest sounding room
4_______
3_______
2_______
1_______ Smallest sounding room

Did the perceived room size affect the vocals positive or negative in the mix?
Give maximum two reasons why the perceive room size hade a positive or negative effect on the vocal sound. If there were no positive or negative affect of the perceived room size on the vocals leave the stimuli blank.

Stimuli A: Positive Negative
Why?

Stimuli B: Positive Negative
Why?

Stimuli C: Positive Negative
Why?
Rank the 5 lead vocal environments from mostly preferred to the least preferred by you in the mix?

Rank the stimuli’s in the order of your liking and write the letter of the stimuli beside the corresponding number on the scale, a 0.5 is perceived as the least preferred and a 10 is the most preferred, a 0.5 must be present as the least preferred and a 10 must be present as the most preferred, no stimuli can’t have the same value as another stimuli.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Most preferred</th>
<th>Least preferred</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td></td>
<td>0,5</td>
</tr>
<tr>
<td>9,5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8,5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7,5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6,5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5,5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4,5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3,5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2,5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1,5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Why is it the most preferred?

Why is it the least preferred?
Why is it the least preferred?


Thank you!